Drugs Explained

Peter L. Bonate, PhD

Drugs Explained

Peter L. Bonate, PhD

Drugs Explained

By Peter L. Bonate, PhD

This work is subject to a Creative Commons CC-BY-ND license.

Subject to such license, all rights reserved.

Copyright © 2026 by Peter L. Bonate

Printed in the United States

ISBN: 979-8-3507-4578-8

This book may not be reproduced in whole or in part, in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system now known or hereafter invented, without written permission from the author.

Dedication

This book is dedicated to my wife and muse, Diana, the great love of my life, who, for some strange reason, 40 years later, never fails to support and encourage me in these Quixotic activities.

This book was written by a human (me) and not by large language models. All materials in the book are my own opinions and do not represent those of my employer.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Preface

I have worked in the pharmaceutical industry most of my life, on numerous products that reached the market and many more that did not. Over the years, I have learned how drugs are developed. And while I may have learned the process, I have noticed that many people working at pharmaceutical companies do not know what we actually do. They may say we develop and make drugs. This is true, but it trivializes the process. What does it mean to develop a drug? What are the activities that are necessary to bring a drug to market? There are whole groups of individuals in companies who don't know. These may include the administrative staff, the legal team, the contracts team, the chemists, and the list goes on. Even among scientists employed by a company, while they may be experts in their particular area, they may not be familiar with the entire drug development process.

Starting around 2013, I began developing a series of lectures at Astellas to explain the big picture of drug development, with an eye towards the drug label, so participants could better understand how to read it and comprehend the activities involved in creating it. This course was named – Drugs 101. This book is an extension of those lectures. Each chapter was a lecture in the course. Let me state up front, I know this book is decidedly US-centric in its approach. I barely mention laws other than the Federal Food, Drug, and Cosmetic Act or other regulatory agencies other than the FDA. This is the "write what you know philosophy." Perhaps in the future, I can expand into other regions and understand how their laws and regulations affect drug development. For example, privacy laws are much stricter in the EU; what impact does this have on drug development?

I've had so much fun writing this book. I've written many books, and this one is a far departure from my usual fare. I've decided to take a more conversational tone to make it more reader-friendly. Drug development is fascinating and full of little tidbits of trivia and stories that are interesting (at least to me). I hope you find these little bits of trivia and history as interesting as I do.

I also wanted to try something different with this book: make it a living document available online. I plan to update it regularly, rather than releasing editions every few years. I also don't plan to sell it to make money. I will, however, make it available on Amazon for those who want a hard copy. I learned a long time ago that if you want to make money, don't write a science book. The only one making money on publishing is the publisher. But I don't write these books to make money anyway. I write them because I have something to say, and it has to come out, much like why a musician writes music or an artist creates their art — not that I am comparing myself to a musician or an artist. Lol.

I've got lots of ideas for the future updates: new stories, reader-requested topics, and new chapters. I want to add a chapter on the sales and marketing of pharmaceuticals, but it will require significant time since I am not an expert in this area. Look for that in the future. I also want to write material on food regulation (Drugs and Food Explained?) and medical devices, perhaps even cosmetics.

I am trying to figure out the best way to disseminate the book. I will be making it available (eventually) on Github and Zenodo. I will also be working on a personal website at peterbonate.com. If there is something you want to see, don't hesitate to write me at peter@peterbonate.com.

I want to thank all those who contributed to this book. Many researchers graciously allowed me to reprint their figures. What would I have done without Wikipedia and Wikimedia Creative Commons for many of the images in the book? Thank you.

I would also like to thank the many reviewers of the book who helped me make it more interesting and relevant. Thank you to my team: Souvik Bhattacharya, Mary Choules, Alexis Hoerter, Hu Huang, Jace Nielsen, Pouye Sedighian, Vijay Siripuram, Junko Toyoshima, and Jared Weddell. You were all a great help in making this more cohesive. I didn't want to write the regulatory chapter, but I listened

 This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

to your advice, and you were right; the book needed it. I want to thank Salaheldin Hamed for his review of that material. His comments made it a lot better overall. I also want to thank my friend, Christiane Collins, who never saw a contraction she liked. Thank you very much for reading this book and giving me your thoughts.

Peter Bonate, March 2026

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table of Contents

What is a Drug? 1
Foods, Drugs, and Cosmetics 1
Natural Does Not Mean Safe: Supplements Explained 4
Small Molecules and Biologics 11
Cell Therapy 18
Gene Therapy 18
Brand-name Drugs vs. Generics 21
The Dose Makes the Poison 24
Summary 25

A Brief History of Drug Discovery 27
Before World War II 29
World War II to 1980 – The Golden Age of Drug Discovery 35
The 1980s to 2000s – Rise of Biotechnology 48
The 2000s to 2020s 58
The Future 67

Regulatory Considerations 71
The Federal Food, Drug, and Cosmetic Act (FFDCA) 71
Public Health Service Act for Biological Drugs 72
Structure of the FDA 73
Interactions with the FDA 74
Types of Applications 76
Drug Approval Pathways 81
Risk Evaluation and Mitigation Strategy (REMS) 84
Post-Marketing Commitments and Requirements 86
The Prescription Drug User Fee Act 86
Over-the-Counter Drugs 86
Orphan Drugs 88
Access Programs 88
Role of Advocacy Groups in Drug Development and Regulation 89
The Package Insert, aka the Product Label, aka Prescribing Information 90
Off-Label Use 95
Drugs of Abuse 99
How Do Drugs Get Their Names? 99
Patents and Exclusivity 102
Non-US Regulatory Authorities 104
Summary 105

Principles of Pharmacokinetics 107
Pharmacokinetic Measures of Exposure 109
The Concept of Clearance 110
The Concept of Volume of Distribution 112
The Concept of Half-life 113
Drug Absorption 115
Bioavailability 119
Drug Distribution 121
Drug Metabolism 125

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drug Elimination ... 136
Kinetics of Multiple Dosing and the Concept of Steady-State ... 140
Linear and Nonlinear Pharmacokinetics ... 144
The Concept of Bioequivalence ... 146
Pharmacokinetic Variability ... 148
Pharmacokinetics of Biologics ... 148
Pharmacokinetics of Cell Therapies ... 152
Flat Dosing or Weight-Based Dosing? ... 154
Summary ... 156

Principles of Pharmacodynamics ... 157
Drug Targets ... 157
Receptor Theory ... 159
Receptor Occupancy and Drug Efficacy ... 167
Enzymes and Transporters ... 170
The Dose-Response Relationship ... 173
Kinetics of Drug Response – Single Dose ... 174
Kinetics of Drug Response – Multiple Dosing ... 181
Time-Dependent Drug Effects ... 182
Chronopharmacokinetics and Chronopharmacodynamics ... 184
And Now For Something Completely Different ... 186
Summary ... 187

Drug Development ... 189
The Process of Drug Discovery to Lead Optimization ... 189
Preclinical Development ... 195
Clinical Development ... 205
Investigator-Initiated Trials ... 229
Registries ... 230
Why Do Drugs Fail? ... 231
Summary ... 233

Efficacy and Safety ... 235
Efficacy ... 235
Safety and Adverse Events ... 246
Benefit-Risk Assessment ... 259
Not All Adverse Events Are Bad ... 260
Summary ... 261

Special Populations ... 263
Introduction ... 263
Pregnancy and Lactation ... 265
Renal and Hepatic Impairment ... 269
The Elderly ... 273
Pediatrics ... 276
Obesity ... 278
Summary ... 282

Drug Interactions ... 285
Example of a Drug-Drug Interaction: The Terfenadine Story ... 285
Example of a Drug-Beverage Interaction: Fexofenadine and Grapefruit Juice ... 288

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Example of a Drug-Supplement Interaction: St. John's Wort 294
Example of a Drug-Smoking Interaction: Tarceva 294
Example of a Drug-Food Interaction: Tykerb 295
Example of a Drug-Protein Interaction: Tocilizumab 297
Example of a Drug-Food Interaction: MAO Inhibitors 298
Example of a Drug-Chemical Interaction: Antacids 300
Example of a Drug-Pharmacodynamic Interaction: Tirzepatide 301
Example of a Drug-Microbiome Interaction? Checkpoint Inhibitors and Proton Pump Inhibitors 302
Example of a Drug-"Weed" Interaction 302
Not all Drug Interactions are Bad 306
How Drug Interactions Are Studied Today 307
How Drug Interactions are Reported on the Label and Medication Guide 309
Summary 313

Personalized Medicine 315
Companion Diagnostics 315
Pharmacogenetics and Pharmacogenomics 319
Case Study in Pharmacogenetics – Drug Metabolism 324
Study in Pharmacogenetics – Pharmacodynamics 329
Why Isn't Pharmacogenetics or Pharmacogenomics Used More Frequently in Clinical Practice? 332
The Future of Personalized Medicine 332
Summary 336

Appendix 339

Appendix 1: More on Estimating Human Doses from Animals and First-in-Man Dose Selection 341

Appendix 2: Design and Interpretation of Clinical Trials 347
Phases of Drug Development Refresher 347
Common Clinical Trial Study Designs 348
Clinical Trial Elements 356
Recent Advances in Clinical Trial Design and Development 361
Statistical Analysis and Reporting of Pivotal Trials 363

References 369

Index 389

About the Author 395

Version Control 397

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 1

What is a Drug?

Drugs are all around us. Got a headache? Take a pill. Got a stomachache? Take another pill. Got a weird rash? Here's a cream for you. Drugs have permeated our culture. We write songs about them – "I want a new drug," "Lucy in the Sky with Diamonds." And they are in the news—the opioid crisis, the EpiPen shortage, mRNA COVID vaccines, and GLP-1 agonists to treat obesity. Athletes are constantly being banned from one sport or another for using performance-enhancing drugs. The legalization of medical marijuana has been rapid in many states. Drugs are *everywhere*. And while the title of this chapter is somewhat glib, defining what a drug 'is' can be difficult.

First, let's take a little test. Go ahead and write down in your own words what a drug is. And then, after reading this chapter, let's see how you did.

Foods, Drugs, and Cosmetics

One way to define a drug is through its legal definition. As defined by the Federal Food, Drug, and Cosmetic Act (FFDCA, Sec. 201; 21 USC 321), a set of laws initially approved in 1938 that gives the Food and Drug Administration (FDA) the authority to oversee the safety of food, drugs, and cosmetics, a drug is defined as [2]:

> *"(A) articles recognized in the official United States Pharmacopoeia, official Homoeopathic Pharmacopoeia of the United States, or official National Formulary, or any supplement to any of them; and (B) articles intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man or other animals; and (C) articles (other than food) intended to affect the structure or any function of the body of man or other animals; and (D) articles intended for use as a component of any article specified in clause (A), (B), or (C). A food or dietary supplement for which a claim, subject to sections 343(r)(1)(B) and 343(r)(3) of this title or sections 343(r)(1)(B) and 343(r)(5)(D) of this title, is made in accordance with the requirements of section 343(r) of this title is solely because the label or the labeling contains such a claim. A food, dietary ingredient, or dietary supplement for which a truthful and not misleading statement is made in accordance with section 343(r)(6) of this title is not a drug under clause (C) solely because the label or the labeling contains such a statement."*

This is too complicated; only a lawyer might understand what it means. But hidden within there is the meat of the definition, and that is clause (B): a drug is an *article intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man or other animals*. On the surface, this may seem an acceptable definition, but it does not cover those drugs that are not used for those reasons. For example, some drugs of abuse, like LSD, have no medical purpose but are still considered drugs.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

A more useful lay definition might be something like:

- a drug is a substance that has a physiological effect (and may have a multitude of effects) when ingested or otherwise introduced into the body,
- is not intended for sustenance, and
- does not create responses that are not already inherent in the body.

But even this might be an outdated definition because what about artificial cells, like chimeric antigen receptor T-cells (CAR-T cells), used to treat cancer? Are those a substance? What even is a 'substance'? Nevertheless, cell therapies are indeed drugs. Pharmacologists might say drugs work through receptors (which will be discussed later), but this does not explain a drug like mannitol, which is used as a diuretic and works not by binding to a receptor, but by increasing plasma osmolality, a non-receptor mechanism.

Properties of Drugs

- Effect depends on the dose administered;
- Can have a multitude of effects;
- Are not taken for sustenance; and
- Modulate or interact with the body's existing physiology; they do not create new responses not inherent in the body.

Drugs must at least have some effect on the body; they turn on, turn off, promote, or inhibit effects that are already present in the body; and are generally distinguished from foods by their lack of sustenance. But how do drugs differ from medicines? Medicines are not recognized in the FFDCA; only drugs are. While they are often used interchangeably and overlap in their definitions, they are different because a medicine is a substance used to prevent or treat the symptoms of a disease for therapeutic benefit. Medicines are a subset of drugs; that is, all medicines are drugs, but not all drugs are medicines. Hence, the FFDCA definition of a drug is more akin to that of a medicine.

The FFDCA distinguishes food from drugs. How do they differ? Food in the FFDCA has a circular definition in that it is defined as:

1. *"Articles used for food or drink for man or other animals,*
2. *chewing gum, and*
3. *articles used for components of any such article."*

This is not a readily digestible explanation (pun intended). Food will be defined as any ingestible substance taken for sustenance to maintain life and growth.

But this begs the question: Can food also be a drug, or vice versa? Legally, no. Food and drugs are distinct in the eyes of the law. How the FDA determines whether a product is a drug or food is determined by the Intended Use Doctrine [3]. This doctrine arose from a 1983 legal case involving Nutrilab. This company extracted a protein from kidney beans that blocked the absorption of starches into the body and marketed it as a weight-loss product known as starch blockers.

The FDA declared starch blockers to be drugs, requiring Nutrilab to file a New Drug Application or remove the product from the market. Nutrilab sued the FDA, claiming that starch blockers, natural products derived from kidney beans, were food, not drugs. The judge in the case found for the FDA because 1.) food, as defined in the FFDCA, is defined as "articles used for food or drink"; articles derived from food are not stated to be food; and 2.) it cannot be said that Nutrilab intended the product to be sold as food; they were marketing them as a treatment

Are These Drugs?

- Is oxygen a drug?
- Can soap be a drug?
- Can dirt be a drug?
- Can ground concrete be a drug?

Check your answers on the next page.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

for being overweight. Nutrilab later appealed the case and lost; as a result, starch blockers are now marketed as supplements.

Based on these court findings, the FDA generally considers objective intent, specifically how the manufacturer promotes and labels the product for its intended use. Typically, when a manufacturer promotes the product for its "taste," "aroma," or "nutritional value," then it is considered a food. Legally, a product generally used as a drug cannot be considered food, and vice versa. In practice, the line is sometimes gray. Some foods can also be drugs. Any food that meets one or more of the FFDCA criteria can be considered food. In addition, foods can also serve as vehicles for drug administration.

The gray zone between food and drugs is also occupied by other foods that have pharmacological effects, the most prevalent being alcohol and coffee. Alcohol certainly has stimulant and depressant effects. Foods like chocolate contain caffeine, which increases arousal and awareness. Coffee is in a similar situation – it has long been regarded as a primary source of caffeine. Other foods, like nutmeg, sweet potatoes, and turkey, can also affect brain chemistry. Then you have drugs that are used as food, like marijuana-laced brownies or alcohol. Some alcoholics look to alcohol to provide some of their caloric intake, and, hence, alcohol acts like a food in this case.

The last item regulated by the FFDCA is cosmetics. As defined by the FFDCA, a cosmetic is:

1. *"Articles intended to be rubbed, poured, sprinkled, or sprayed on, introduced into, or otherwise applied to the human body or any part thereof for cleansing, beautifying, promoting attractiveness, or altering the appearance, and*
2. *articles intended for use as a component of any such articles; except that such term shall not include soap".*

Articles in this group include lipsticks, perfumes, deodorants, hair color, skin moisturizers, and makeup for the eyes and face. Soaps are expressly excluded from this list (see the blue box on "Are

Are These Drugs?

- Is oxygen a drug? Yes. Although oxygen is essential for life, the percentage of oxygen in breathable air is relatively small, at approximately 20%. Hyperbaric chambers increase the oxygen content of the air to 100% and help treat conditions such as gangrene, burns, and necrotizing soft-tissue infections. Oxygen can also improve wound healing.
- Can soap be a drug? It depends. The soaps of today are not the soaps of your grandparents' generation. Today's soaps are complex products of chemicals, synthetic detergents, and, in some cases, antimicrobial agents. Although soaps were explicitly excluded from the original 1938 FFDCA, they are now subject to review for their ingredients and purpose. If the soap is used to cure, mitigate, treat, or prevent disease, it may be classified as a drug. Suppose the soap consists of detergents or is primarily composed of alkali salts of fatty acids and is intended not only for cleansing but also for other cosmetic purposes. In that case, it may be regulated as a cosmetic.
- Can dirt be a drug? No, but it may be a dietary supplement. Dirt is a mixture of minerals and organic matter that contains microorganisms. It has been estimated that we may consume as much as 6 lbs. of dirt a year in our food. The organisms in dirt may act like a probiotic, increasing the health of our gastrointestinal tract, and the minerals may serve as a supplement. However, that doesn't mean we should rush out and start eating dirt, as some dirt also contains pesticides and toxins. If you're going to eat dirt, be aware of what you're consuming first.
- Can ground concrete be a drug? No. I threw this in the mix to see if you were paying attention.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 1 Facts About Cosmetics			
Cosmetics have their own set of Good Manufacturing Practices.	For instance, some products regulated as cosmetics in Europe are regulated as drugs in the United States (US), e.g., sunscreens.	Cosmetics imported into the US, including both ingredients and finished products, must meet the same safety and labeling criteria as those manufactured domestically.	FDA does not pre-approve cosmetic products or ingredients, with the important exception of color additives.
Cosmetic firms are responsible for marketing safe, properly labeled products without prohibited ingredients and adhere to restricted ingredient guidelines.	Color additives must be certified for purity in FDA labs if they are to be used legally in a product marketed in the United States.	FFDCA does not specify any particular testing regimens for cosmetic products or ingredients.	The cosmetic company is responsible for substantiating the safety of its products and ingredients before marketing them to the public.

These Drugs?" for clarification). Can cosmetics be drugs? Yes. Sometimes, you may hear the term 'cosmeceutical' used to describe a product that combines the properties of both cosmetics and drugs, such as a shampoo that has both cleaning and antibacterial properties, or retinol-containing anti-aging creams. However, under the law, the FFDCA does not recognize the term. The FDA uses the Intended Use Doctrine to determine whether a cosmetic is a cosmetic, a drug, or, in some cases, a medical device. Cosmetics do not require FDA approval before they can be marketed and are the least regulated among items under the FFDCA. However, in 2022, a significant revision to the FFDCA was issued, enhancing the FDA's regulatory jurisdiction and oversight of cosmetics [4].

Natural Does Not Mean Safe: Supplements Explained

There is a movement that began in the 1970s, when Linus Pauling, a Nobel Prize-winning chemist and Nobel Peace Prize winner, proposed megadoses of Vitamin C to prevent the common cold, and later, with Ewan Cameron, a Scottish physician, proposed high-dose Vitamin C as a cancer therapy. Pauling was a major proponent of orthomolecular medicine, which proposes that disease can be prevented by creating an optimal molecular environment within the body through the use of vitamins and micronutrients. Pauling's stance tainted his Nobel Prize win, as numerous studies have failed to demonstrate that Vitamin C could reduce the incidence of common colds [5]. Nevertheless, the idea took hold in the public consciousness and cannot be let go. This idea that 'natural is healthy' is firmly rooted. Celebrities like Gwyneth Paltrow, Madonna, and Adele continue to hold influence over the public, advocating mega-doses of Vitamin C and vitamin infusions for better health.

The concept of vitamin supplements has expanded to include other natural products. For example, turmeric has been proposed as an anti-inflammatory agent, echinacea has been used to treat colds and boost immune function, and garlic has been used to enhance cardiovascular health and immune function. The problem is that some of these natural remedies are, in fact, therapeutic. For example,

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 2 Similarities and Differences Between Drugs and Supplements		
Property	**Supplements**	**Drugs**
Regulatory Basis	DSHESA 1994	FFDCA
Premarketing Approval	Not required	Required
Content Claims	Allowed	Allowed
Structure/Function	Allowed	Not used; drugs make direct therapeutic claims
Disease Treatment Claims	Not allowed unless approved by the FDA	Allowed
Disclaimers	Yes for all structure/function claims	Claims are FDA-approved
Evidence of efficacy required	No	Yes
Adequate safety	Yes	Yes
Follow GMPs	Yes	Yes

folic acid has been shown to prevent neural tube defects during pregnancy, zinc can reduce the duration of colds by about 2 days, and peppermint oil is a useful short-term remedy for irritable bowel syndrome. However, not all supplements can improve health and treat illness. Many supplements lack medical value. For example, vitamin E has been shown not to affect cardiovascular outcomes, despite advertisements that say otherwise. And some supplements are even dangerous. Ephedra was an herbal supplement that promoted weight loss and increased energy; it was banned in 2004 by the FDA because of reported high blood pressure, heart attacks, stroke, seizures, and death associated with its use. The message here is that 'natural' does not mean 'safe'.

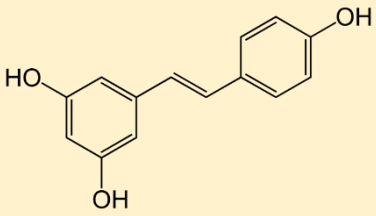
Due to the increasing use of supplements in the 1980s and 1990s and the cries of "FDA overreach" in attempts to regulate these products, Congress modified the FFDCA through the Dietary Supplement Health and Education Act (DSHEA) in 1994, creating a new category called dietary supplements. Hereafter, supplements will be treated differently from drugs (Table 2). As defined, a dietary supplement:

> *"means a product (other than tobacco) intended to supplement the diet that bears or contains one or more of the following dietary ingredients:*
>
> *(A) a vitamin;*
>
> *(B) a mineral;*
>
> *(C) an herb or other botanical;*
>
> *(D) an amino acid;*
>
> *(E) a dietary substance for use by man to supplement the diet by increasing the total dietary intake or*
>
> *(F) a concentrate, metabolite, constituent, extract, or combination of any ingredient described in clause (A), (B), (C), (D), or (E).*
>
> *Dietary supplements are further defined as products labeled as dietary supplements and are not represented for use as conventional food or as a sole item of a meal or diet. Supplements can be marketed for ingestion in a variety of dosage forms, including capsule, powder, soft gel, gelcap, tablet, liquid, or, indeed, any other form so long as they are not represented as conventional foods or as sole items of a meal or of the diet (FDCA, as amended, § 402)."*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

What Went Wrong with Resveratrol?

In 2003, newspapers around the world announced that the Parisian paradox had been solved [6]. Parisians live as long or longer than others, despite consuming a diet high in fatty foods that are generally considered unhealthy for the heart. Scientists at Harvard reported on a protein called SIR2 that could increase yeast lifespan by an astounding 70%. Humans don't have SIR2, but they have a similar protein called SIRT1. Scientists at Harvard screened a library of compounds for their ability to activate SIRT1 and identified 17 compounds. The most potent of these was resveratrol, which is found in high concentrations in some grapes and red wines. Later studies demonstrated that resveratrol is an antioxidant with protective properties that can slow aging in human cells. After the initial announcement, resveratrol received attention from "60 Minutes" and "The Oprah Winfrey Show." Many made the association that, since many Parisians drink a lot of red wine, this was the solution to the Parisian paradox. The Harvard Scientists quickly spun out a company called Sirtris to develop resveratrol as a drug, which GlaxoSmithKline bought for $720 million in 2008. Additionally, many companies have begun marketing resveratrol as a dietary supplement. To this day, resveratrol is not a marketed drug, despite its ongoing sales as a supplement. What happened?



Early analysis of resveratrol pharmacokinetics showed < 1% was absorbed orally, and plasma concentrations were very low in humans. The concentrations needed to prolong cell lifespan in a test tube would require very large doses (in the gram range) in humans to achieve similar values. SRT501 was developed as a microionized formulation, aiming to improve oral absorption. Even with the new formulation, high doses would be required for pharmacologic activity. Phase II studies in patients with multiple myeloma showed that the addition of 5 g daily SRT501 to bortezomib treatment did not show evidence of improved efficacy, but there were 5 cases of severe renal toxicity, leading the patients to be withdrawn from the study [7]. Following numerous failed studies and the poor pharmacokinetic properties of reservatrol, GSK discontinued all further development of reservatrol in 2011. Today, however, researchers continue to search for the Holy Grail – a suitable formulation and dosing conditions that can enable resveratrol to be administered at relatively low doses without adverse effects. What is interesting about this is that many supplements, which do not have any special formulation to improve bioavailability, recommend 150 to 500 mg per day for a healthy heart, 500 to 1000 mg per day for cancer prevention, and 100 to 300 mg per day for improved cognitive function.

Studies continue to report clinical benefits of resveratrol, but the consensus on its medicinal value remains conflicting. For every study showing benefit, there is another that does not. Many social media influencers who advocate for its use cherry-pick articles to support their claims.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

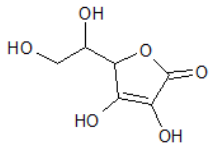
Based on this definition, most dietary supplements contain one or more of the following: vitamins, minerals, herbs, or other botanical parts (such as stems, leaves, or bark), amino acids, or protein powders and shakes. The Nutrilab case falls under (F), which is the extract of the kidney bean.

Dietary supplements are presumed safe, but 'safe' is a relative term. Water-soluble vitamins, such as Vitamin B and C, are relatively safe in high doses because they are excreted in the urine. However, toxicity may occur after taking too high a dose. Greater caution must be exercised with the fat-soluble vitamins A, D, E, and K. Too much of the fat-soluble vitamins can accumulate in the body and lead to:

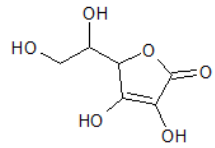
- Vitamin A: nausea, increased intracranial pressure, and possibly death;
- Vitamin D: weight loss, appetite suppression, and cardiac arrhythmias; and
- Vitamin E: changes in blood clotting leading to bleeding or hemorrhagic stroke.

Vitamin K is relatively safe, but overdosing can occur quickly for Vitamins A, D, and E. For example, hypervitaminosis A can occur with a single dose of 200 mg, which is 10 times the daily recommended allowance.

Here is the structure of Vitamin C made in an orange and Vitamin C made in the laboratory by chemical synthesis:



Vitamin C from an Orange



Vitamin C Made in the Lab

Notice the difference? No? That's because there isn't any. There is a belief among many that natural-made products are "better" than synthetic products made in the laboratory. For all the vitamins, there is absolutely no difference between synthetically created vitamins and naturally made vitamins. In fact, companies like Nature Made make their vitamins using microbial fermentation in large-scale bioreactors. For herbal products, such as St. John's wort, the plant is grown in large-scale fields, cultivated, and the active ingredient is extracted and purified before being formulated into a product, typically a tablet or capsule. Sounds good, doesn't it? In theory, maybe, but in practice, the product quality of herbal products is not very good.

One study examining the DNA of 5,957 commercial herbal products sold in 37 countries showed that 27% of products were adulterated, which means they contained undeclared contaminants, substitutes, or fillers, or none of the labeled herbal product [8]. The countries with the highest levels of adulteration were Australia (79%) and South America (67%). Europe and North America were 47% and 33%, respectively. This study is not unique. Booker et al. [9, 10] compared the product quality of 47 commercially available St. John's work products and 40 rosea products. 36% of St. John's wort products were adulterated. 20% of rosea products did not even contain rosea. 80% of the remaining products had lower content than reported on the packaging.

And if you think things are better with vitamins and other supplements, you would be wrong. Vitamins and supplements are often mislabeled and frequently do not contain the reported content. For example, one analysis of 8 Coenzyme Q10 (CoQ10) products sold on Amazon reported that four of these products had deceptive labeling. For example, the label reported the content as "400 mg/6%", but it actually contained 24 mg of CoQ10. Others reported significantly less content than the label. One brand had a labeled amount of 200 mg but actually contained only 11 mg. Another had 21 mg but listed 400 mg on the label. Of the eight brands tested, only one had the actual amount on the label [11].

But how can supplement manufacturers get away with this? They can because supplements are not held to the same product quality standards as pharmaceutical manufacturers. The latter have strict requirements for chemistry, manufacturing, and control. Supplement manufacturers do not.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

More Does Not Always Mean Better with Drugs and Supplements

There is an attitude in society regarding drugs that "more is better." Many have an attitude like "I have a headache, and the box recommends two pills every 6 hours. I should take four pills every 6 hours to get a greater benefit. If two is good, four will be great!" While this may work for donuts and coffee, this is a potentially dangerous attitude to have about drugs and supplements, and there are many reasons to be aware of:

1. All drugs have a therapeutic window, which is a range of concentrations in the body where the drug is effective without being toxic. Taking a larger dose could move you out of the therapeutic window into an unsafe area where adverse events could occur. Acetaminophen (Tylenol) will be discussed extensively in this book because it is so widely used. The recommended dose for Tylenol in adults is 650 to 1000 mg every 4 hours, but no more than 4 g per day. Single doses in the 7 to 10 g range can lead to severe liver damage and, maybe even, death. That's not much of a difference between the recommended dose and unsafe dose, and considering how many products contain acetaminophen, this leads to many accidental overdoses around the world.
2. Drug concentrations may not increase linearly. If you double the dose, the concentration in the body may not necessarily double. Sometimes you can have much higher concentrations than normal because the enzymes that metabolize the drug and remove it from the body are saturated. Sometimes you saturate the proteins that help the drug be absorbed into the body, limiting how much of the drug is actually absorbed and leading to increased excretion of the drug into the feces.
3. Often, what you see is not an increase in efficacy, but an increase in adverse events. Most recommended drug doses are already titrated to achieve the "optimal" effectiveness with the least side effects. By increasing the dose, there is little room for effectiveness to increase, but lots of room for both the severity and frequency of adverse events to increase.
4. Increased doses could lead to tolerance or resistance to the drug, reducing its effectiveness. For example, increased opioid frequency for pain can lead to tolerance and reduced effectiveness. Increased antibiotic use can lead to resistance in the organisms the drugs target, reducing or sometimes eliminating the drugs' effectiveness.

Many supplements are sold as mega-doses. Multivitamins, for example. For some supplements, the therapeutic window is large, and mega-doses are relatively safe. But for some supplements, it is not safe. Too much iron can lead to gastrointestinal pain and bleeding. Mega-doses of B-vitamins (more than 200 mg Vitamin B6 per day) can lead to nerve damage, and long-term administration has been linked to a 30 to 40% increased risk of lung cancer in males [12]. It has been estimated that there are more than 23,000 Emergency Room visits every year because of supplement overdoses, 2000 of which require hospitalization [13]. The bottom line is to be cautious when increasing the recommended dose for a drug, and to be especially careful with supplements in general.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 3 Differences Between Small Molecules and Biologic Drugs	
Biologics	**Chemical Drugs**
Made by living cells	Made by chemical synthesis
Large molecular weight and size	Small molecular weight and size
The structure is complex and heterogeneous	Well-defined structure
Difficult to structurally characterize analytically	Easy to structurally characterize analytically
Sensitive to environmental conditions	Relatively stable
Usually injected	Usually taken orally

Supplements must be within 90 to 110% of the amount reported on the label and demonstrate purity, free of contaminants. In a market with over 4,000 unique products and more than 75,000 products on shelves nationwide, the FDA cannot possibly ensure product quality for all supplements sold in the US. It is generally recognized that 20% of supplements are adulterated in some form [14], and this number may be on the low side.

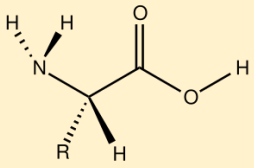
Supplement manufacturers have different responsibilities than drug manufacturers. Supplement manufacturers must ensure that their products are safe and that their product labels are truthful and not misleading. Their safety aspect differs from that of drugs, where manufacturers must demonstrate both safety and efficacy. Supplement manufacturers are not required to demonstrate that their products are efficacious, and they cannot make the same claims as drug manufacturers.

On their labels, supplements are allowed to make content claims, e.g., "this product contains 100 mg Vitamin C," or "High in Vitamin C," or "100% of the daily requirement for Vitamin C" (per 21 CFR §101.93(f). Supplements may also make structural/functional claims about how they affect the structure or function of the body, e.g., "Calcium builds strong bones" and "Supports Immune Health." These claims do not require FDA approval, but they must be truthful and cannot be misleading. Specific health-related claims, like "Calcium may reduce the risk of osteoporosis", are allowable but must be approved by the FDA. What a supplement cannot claim is that it can diagnose, treat, cure, or prevent disease – these claims are reserved for drugs. Hence, on most supplement labels, there will be a statement that reads "This product is not intended to diagnose, treat, cure, or prevent disease."

Somewhere between drugs and food, and somewhat related to supplements, is a whole new category of products called nutraceuticals. First coined in 1989, nutraceuticals are defined as a "food (or part of a food) that provides medical or health benefits, including the prevention and treatment of a disease." There are two classes of nutraceuticals: dietary supplements and functional foods. Dietary or nutritional supplements are designed to supplement and provide nutrients that are not obtained from a standard diet. Examples include glucosamine and chondroitin for joint health, coenzyme Q10 for heart health, fish oil to improve lipid profiles, and resveratrol for cardiovascular health. Functional foods are enriched or fortified foods intended to be consumed in their natural state. Examples include the addition of Vitamin D to create fortified milk, weight-loss bars, and probiotic yogurts. Nutraceuticals are not yet recognized as a separate category in the FFDCA and, as such, are regulated depending on how they are marketed and used.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Proteins

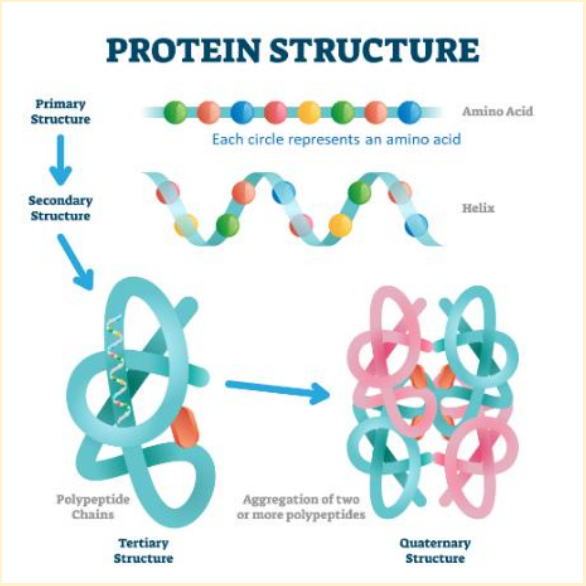


Amino acids are organic compounds with an amino group (-NH_2) at one end and a carboxylic acid (-COOH) at the other end, both of which share a common carbon atom. Attached to the central carbon is a hydrogen (-H) and a unique group (-R) that determines the specific amino acid. There are 20 amino acids in human proteins, which are considered the building blocks of life.

Proteins are essential for life; they are composed of many amino acids, all strung together end-to-end to form a long chain. Whereas peptides are short amino acid chains, ranging from 2 to 50 amino acids in length, with a molecular weight of 6000 Daltons or less, proteins can be much longer. The largest protein in humans, called titin, consists of over 34,000 amino acids and has a molecular weight of ~4 million. Proteins have many functions, such as providing structural support and tissue integrity (keratin and collagen), speeding up chemical reactions (enzymes), being involved with movement (actin and myosin in muscles), providing immune functions (monoclonal antibodies and cytokines), providing transport (hemoglobin), and acting as hormones.

Proteins as drugs have a long history. Early pharmaceutical proteins were extracted from animal organs. They were animal in origin and highly immunogenic. Efficacy did not last long. Today, proteins are produced in giant bioreactor vats by microorganisms. Bioreactor-derived proteins are entirely human, but may still be immunogenic. Examples of protein drugs include insulin, growth hormone, and erythropoietin.

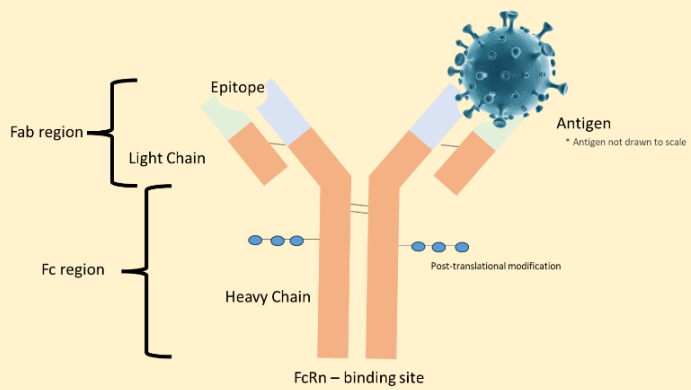
Proteins are characterized by their primary, secondary, tertiary, and quaternary structures. A protein's primary structure is its amino acid sequence. The secondary structure comprises local regions, such as helices and extended sheets, stabilized by hydrogen bonds between amino acid residues. The tertiary structure refers to how a protein folds upon itself and is part of its three-dimensional structure. In contrast, the quaternary structure occurs when multiple protein units combine to form a single, larger entity.



Proteins are also characterized by post-translational modifications (PTMs), chemical alterations that occur after a protein is synthesized. These may include phosphorylation, glycosylation, or acetylation. PTMs modulate a protein's activity and can help direct a protein to a specific location in the body. They may also be responsible for the body forming antibodies against proteins produced in bioreactors. The antibodies may limit the efficacy of the protein-based drug and increase its safety risk.

Figures courtesy of Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Cartoon of a Monoclonal Antibody

Fab region is the binding domain for antigens, confers specificity, and is responsible for therapeutic efficacy

Post-translational modifications are largely responsible for the antigenicity of monoclonal antibodies

FcRn region protects the antibody from degradation and is responsible for antibody recycling, thereby prolonging its half-life in the body

Small Molecules and Biologics

Small-molecule drugs have well-defined chemical structures with molecular weights generally < 1000 Daltons and are chemically synthesized. The majority of drugs approved today are small molecules, and for decades, they have been the major type of drug developed. Examples of small molecules include aspirin, acetaminophen, ibuprofen, and warfarin.

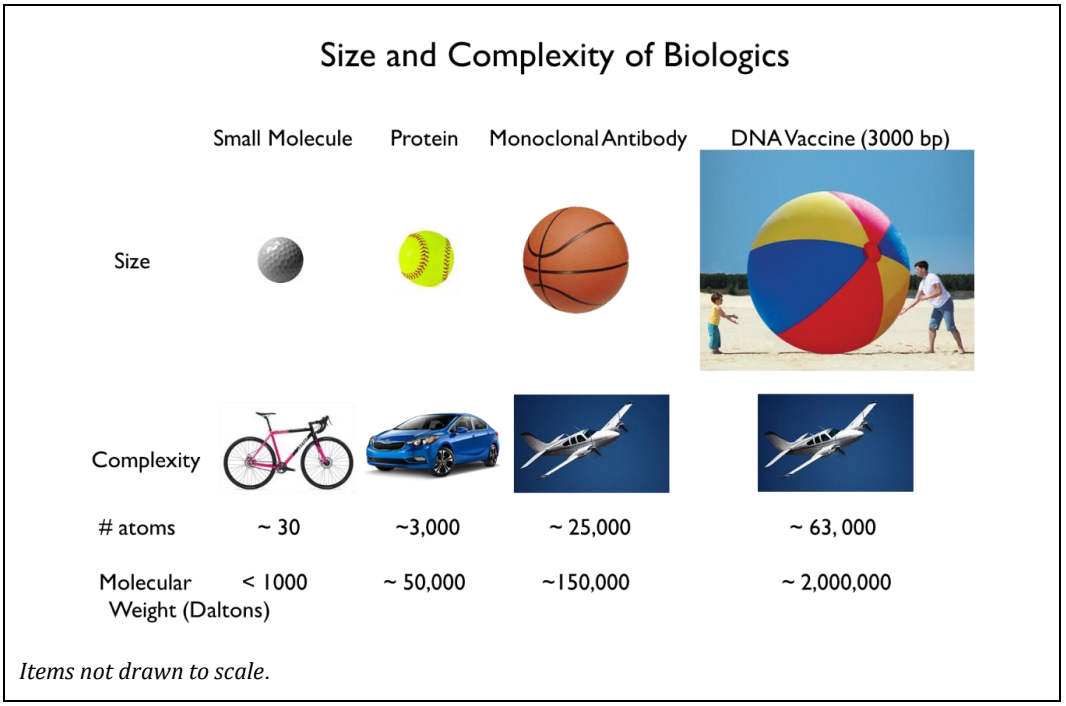
The FDA defines a biological drug as vaccines, blood and blood components, allergenics (such as allergy shots), somatic cells, gene therapy, cell therapies, and recombinant therapeutic proteins. Examples of biological drugs include proteins such as insulin, growth hormone, and interferon; monoclonal antibodies such as trastuzumab (Herceptin); enzymes such as imiglucerase (Cerezyme); antisense oligonucleotides such as fomiversen (Vitravene); vaccines; cell therapies such as tisagenlecleucel (Kymriah); and gene therapy. Most biologics differ from small molecules in that they require injection, either intravenous or subcutaneous, for their administration.

Biological drugs, or biologics as they are often called, are biological macromolecules that are orders of magnitude larger in size and complexity than small molecule drugs (Table 3). Biologics cannot be synthesized directly, unlike small molecules, and are instead produced by living cells — either microorganisms, plants, or animal cells — hence their name. Production of biologics is more complex than that of small molecules, and they cannot simply be made like small molecules. As such, their exact structure and purity cannot always be determined by standard laboratory testing methods. Furthermore, a significant safety concern with biologics is their potential to trigger an immune response, which can be life-threatening.

Monoclonal antibodies represent the dominant class of biologics. As of 2024, roughly 150 have been approved, representing about 60-70% of total biologics sales. The highest-selling drug between 2020 and 2025 was a monoclonal antibody (MAb); Keytruda (pembrolizumab) posted sales of almost $25 billion in 2023 and $30 million in 2024. As such, it is worth discussing what they are.

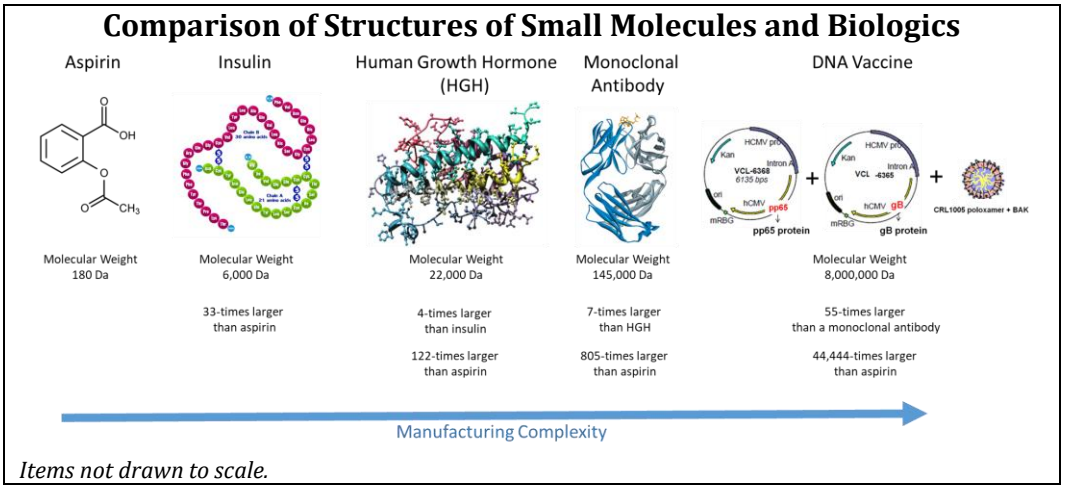
Antibodies are produced by B cells, a type of white blood cell, in response to antigens, which are substances the body recognizes as foreign, such as parts of a bacterium or a virus. Antibodies themselves are large proteins, about 150,000 Daltons, consisting of two heavy chains and two light chains held together by disulfide bonds. At one end is the epitope-binding domain, where the antigen

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



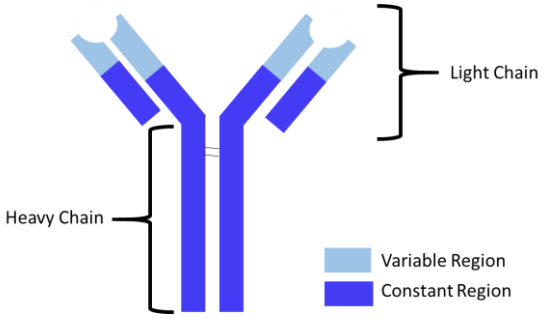
binds to the antibody, and at the other is the FcRn-binding domain, which helps extend the antibody's lifespan in the body. On the surface of the antibody are post-translational modifications, which are chemical changes that occur to a protein after it has been synthesized. These include adding sugar residues to the surface or modifying the antibody with methyl, acetyl, or phosphoryl groups. These changes are important because they are not only essential to the antibody's normal function but also help the body recognize the antibody as part of the person, not a foreign protein.

What normally happens in the body in response to a first-time antigenic challenge is that the immune system detects the antigen, a naïve B cell binds it via its B-cell receptor, and it is internalized into the cell. With the help of T-cells, another type of white blood cell, an army of B-cells is produced against the antigen, which differentiates into two types: plasma cells, which are antibody-producing factories, and memory cells, which remember the antigen for the next time it is encountered by the body, allowing a greater immune response to be achieved. This is a slow process, which takes weeks to months.

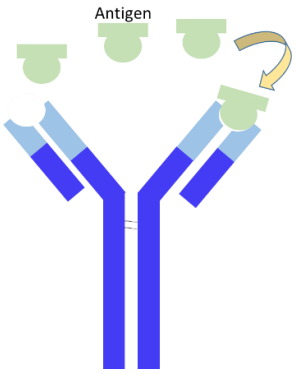


Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

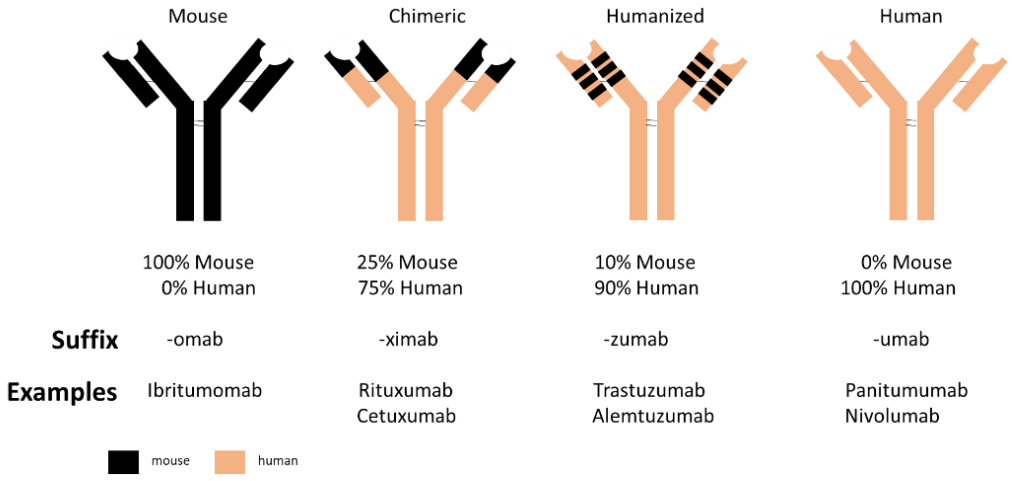
More on Monoclonal Antibodies



Antibodies are Y-shaped proteins consisting of two heavy chains and two light chains. The variable region is said to be the Fab region (fragment antigen binding), and the constant regions are said to be the Fc domains (fragment crystallizable).



Antigens bind to the Fab portion of the antibody. The epitope is the part of an antigen that binds to an antibody.



Early antibodies were either mouse or chimeric, meaning they were either entirely mouse-derived or partly derived from a mouse. These antibodies were highly immunogenic, prompting the body to mount an immune response that led to the formation of anti-drug antibodies that could either increase, decrease, or have no effect on the drug's efficacy. Today, monoclonal antibodies are entirely human, with no murine components.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

For Biologic Drugs, The Process is the Product

Differences in the manufacturing process can result in distinct products, such as Myozyme and Lumizyme.

Monoclonal antibodies are laboratory-produced antibodies produced for a specific epitope by a single B-cell clone. They are lab-made versions of natural antibodies. They differ from the usual antibodies produced by humans in that they are monoclonal (produced by a single clone); human antibodies are polyclonal, produced by multiple B cells. The post-translational modifications observed in normal human antibodies also occur in monoclonal antibodies; however, these modifications can lead to a new type of antibody, known as an antidrug-antibody. Monoclonal antibodies act in the body by a multitude of mechanisms. For example:

- **Binding targets:** Adalimumab (Humira) binds TNF-α in the bloodstream, reducing inflammation;
- **Blocking targets:** Trastuzumab (Herceptin) blocks the HER2-receptor on the cell membrane, stopping oncogenic growth signaling in cancer cells and preventing tumor growth;
- **Cell killing:** Rituximab (Rituxan) binds CD20 on B-cells, which triggers complement-dependent cytotoxicity and antibody-dependent cell-mediated cytotoxicity (ADCC);
- **Agonist:** Urelumab binds the human CD137 receptor and activates its pathway and provides a costimulatory signal to T-cells, boosting the ADCC of NK cells; and
- **Targeted drug delivery:** Antibody-drug conjugates, like Padcev (enfortumab vedotin), are monoclonal antibodies that carry cytotoxic drugs to cells, which are internalized, and then cause cell death.

MAbs have proven so successful that they have led to a host of derivatives, such as bispecific antibodies, antibody fragments, antibodies labeled with radioactive atoms, and Fc fusion proteins, which are antibodies that carry functional proteins, such as receptors, enzymes, or peptides.

For many years, it was said that the "product is the process" for a biological drug. In other words, how the biologic is manufactured is central to how the drug behaves inside the body. This is because macromolecules produced in biological systems, such as bioreactors, are sensitive to changes in the manufacturing process. A change in reactants within the bioreactor, or even in the bioreactor size, can result in a product with different characteristics, potentially leading to variations in the properties of the resulting drugs.

Take, for example, the case of Myozyme and Lumizyme. Pompe disease is a rare orphan disease that affects about 5,000 to 10,000 people worldwide. Pompe patients have an autosomal recessive defect characterized by the absence or deficiency of the alpha-glucosidase enzyme, which breaks down carbohydrates into glucose. This deficiency leads to the accumulation of glycogen in tissues and organs, particularly in muscle. Alglucosidase alfa (AGA) is an enzyme grown in a bioreactor and used to treat Pompe patients by exogenously replacing the defective enzyme. Early clinical trials with AGA were conducted using a product produced in a 160-liter (L) bioreactor. In 2006, AGA was approved by the FDA and EMA and marketed as Myozyme in both regions.

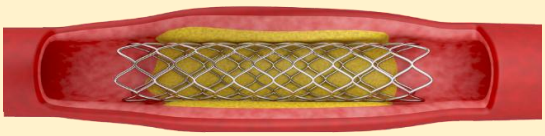
This size bioreactor soon proved insufficient to meet demand, and the production was switched to a 2000 L bioreactor. However, the FDA concluded that the AGA produced by the two bioreactors differed chemically and pharmacokinetically. They were not considered the "same" product, so a switch in bioreactors was not possible. The manufacturer then filed for approval as a new drug product produced by the 2000 L bioreactor, which was granted in 2009, but was marketed under the name Lumizyme. Hence, two versions of AGA were being marketed in the United States (US) by the same company, but produced by different-sized bioreactors. To make matters more confusing, the EMA accepted that the Myozyme created in the 2000 L bioreactor was the same as the Myozyme created in the smaller bioreactor. Hence, in Europe, Myozyme was being sold as the 2000 L bioreactor product, and Lumizyme was not available. Today, Lumizyme is produced in a bioreactor of even larger capacity, 4000 L.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

When is a Drug Not a Drug?

When is a drug not a drug? When it is a combination of a drug and a device. Examples include: the EpiPen, which is sold as epinephrine inside a self-injecting needle unit (auto-injector); drug-eluting stents, which are small mesh tubes designed to keep open blood vessels and slowly elute drug into its local environment; and drug infused implants, which can release a drug into the circulation slowly over weeks, months, or sometimes years, like Exparel (bupivacaine) for non-opioid pain management. Are these drugs? Yes. Are they devices? Yes. But they are considered a combination medical device, containing drugs, and are also regarded as a medical device. The legal definition of a combination device builds on the legal definition of a drug:

> *An instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other such related article, including any component parts or accessories, that contain a recognized and approved drug that is intended to diagnose, prevent, treat, or cure disease in man or animals, that is intended to affect the structure or function of the body, and does not achieve its primary purpose through chemical action alone within the body.*



Sometimes, the distinction between what constitutes a medical device and a drug is blurred. The choice between the two has significant implications, as each has different regulations and approval pathways. For example, is a fentanyl popsicle a drug or a device? Originally approved as a drug, today it might be considered a device. Medical devices that contain new drug entities will need to obtain approval for both the drug and the device aspects.

When modern biologic drugs, particularly monoclonal antibodies and other proteins, started to be developed in the 1970s and 1980s, this presented a regulatory challenge for the FDA because the FFDCA was primarily written for small molecules. However, the Public Health Services Act (PHSA) was an early law to regulate biological products for safety, purity, and potency. In 1902, Congress passed the Biologics Control Act, also called the Virus-Toxin Act, the first US law regulating biological products, like vaccines and antitoxins, after 13 people died when a contaminated diphtheria antitoxin was released to the public. In 1944, the PHSA was enacted, establishing the Public Health Service and placing the regulation of biologics— primarily vaccines, antitoxins, blood, and blood components and their derivatives — under the control of the National Institutes of Health. Later, in 1972, the authority for biologics was transferred from the NIH to the Bureau of Biologics in the FDA, which was renamed the Center for Biologics Evaluation and Research (CBER) in 1988.

When Orthoclone OKT-3 was approved in 1986 as the first therapeutic monoclonal antibody by CBER within the FDA, it established the regulatory precedent of approving the drug under the PHS Act §351, not the FFDCA. Monoclonal antibodies were subsequently regulated under the PHSA by interpretation rather than by explicit authority in the act. The Food, Drug, and Modernization Act (FDAMA) of 1997 strengthened the FDA's biologics authority, but did not alter the PHSA's definition of a biologic. Later, in 2003, CBER transferred regulatory control of monoclonal antibodies to the Center for Drug Evaluation and Research (CDER), the FDA arm that has regulated small molecules, where they are today.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

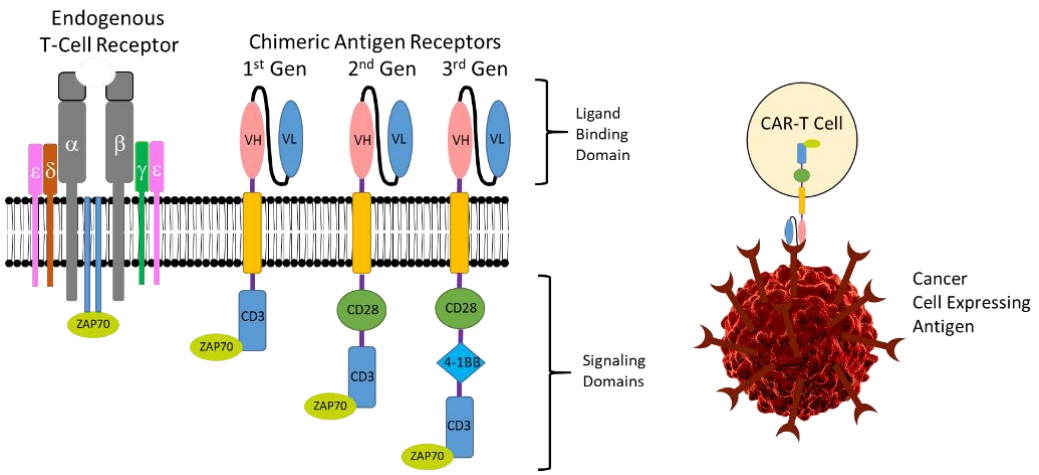


Figure 1. Comparison between endogenous T-cell receptors and 1st, 2nd, and 3rd generation chimeric antigen receptors (CAR) (left). Costimulatory domains have been added to improve their ability to produce more T-cells after infusion and increase their survivability in the circulation. Binding of a tumor cell expressing the binding antigen to a CAR-T cell having the complementary ligand binding domain (right).

In 2010, the Biologics Price Control and Innovation Act, part of the Affordable Care Act ("Obamacare"), officially amended the PHSA to define "biological products" as including proteins, thereby ending the practice of approving MAbs and other proteins on an interpretive basis. At the same time, the biosimilar approval pathway was established for generic biologics. This is not to say that the FFDCA does not regulate biologics – it does. Biologics development, including clinical trials and all manufacturing, falls under the FFDCA, as does labeling and post-marketing safety reporting.

Clusters of Differentiation (CDs) are a method for identifying and defining cell-surface markers on cells and tumors. '+' indicates the CD is expressed on the cell; '-' indicates the CD is not expressed on the cell. For example, stem cells are identified as CD34+CD31-. Immunophenotyping is the process of identifying a cell based on what molecules are present on its cell surface. An example of their use is CD4+ as a marker for helper T-cells and CD8+ as a marker for cytotoxic T-cells. Shown in the figure are the CD markers for different classes of white blood cells.

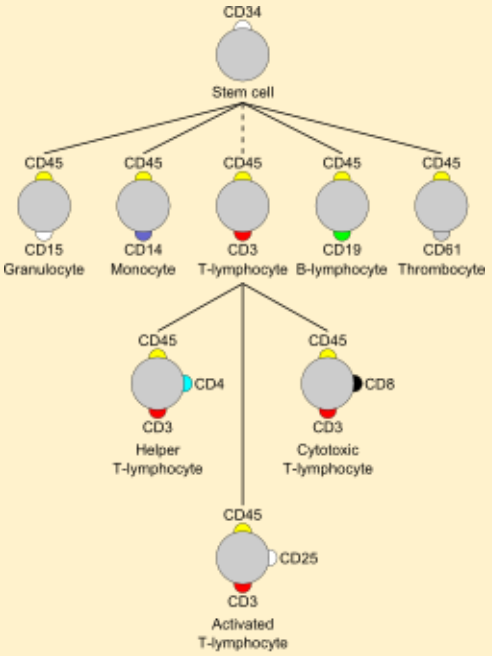
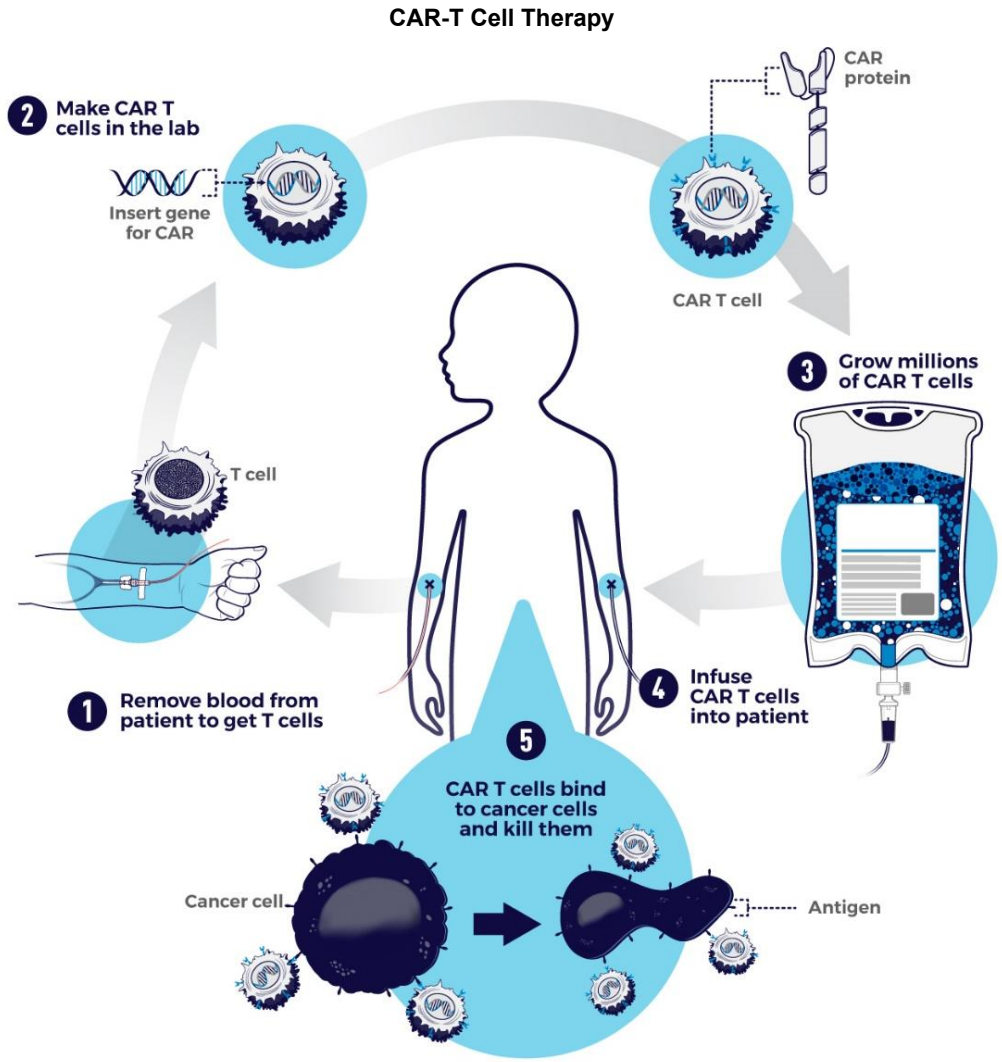


Figure courtesy of Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



CAR-T cell therapy is a type of treatment in which a patient's T-cells are genetically engineered in the laboratory to bind to specific proteins (antigens) on cancer cells and kill them. (1) A patient's T-cells are removed from their blood, and then (2) the gene for a receptor called a chimeric antigen receptor (CAR) is inserted into the T-cells in the laboratory. The gene encodes the engineered CAR protein, which is expressed on the surface of the patient's T-cells, thereby creating CAR-T cells. (3) Millions of CAR-T cells are grown in the laboratory. (4) They are then given to the patient by intravenous infusion. (5) The CAR-T cells bind to the antigens on the cancer cells and kill them.

Figure courtesy of Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Cell Therapy

Another type of biologic drug is cell therapy, the administration of live cells into patients. These cells may be the patient's own cells or universal cells, such as stem cells, and there are many different types. Adoptive cell therapy is the stuff of science fiction. Immune cells from cancer patients, typically T cells, are extracted from the body, isolated, and then expanded ex vivo (outside the body) or modified by adding new receptors. The cells are then reinfused back into patients. In 1986, Steven Rosenberg and colleagues at the US National Cancer Institute harvested tumors from mice, removed T-cells from the tumor microenvironment, known as tumor-infiltrating lymphocytes (TILs), and expanded their numbers ex vivo. They then re-administered the expanded TILs back to the mice. When given in combination with traditional chemotherapy and IL-2, 100% of mice with colon cancer were cured, and 50% were cured of lung metastases [15]. Later, Rosenberg and colleagues administered autologous activated killer cells and IL-2 to patients with advanced cancer, all of whom failed prior therapies. Cancer regression was observed in 11 of 25 patients, with complete tumor shrinkage in one patient for 10 months. While promising, this approach was too toxic for routine therapeutic use.

Enter chimeric antigen receptor T-cells (CAR-T), which are T-cells with chimeric T-cell receptors added to the cell's surface to recognize tumor cells [16]. These cells use antibody fragments to recognize tumor cells. The first chimeric T-cell receptor was developed in 1993 by mixing an antibody-like binding domain to the body of a T-cell receptor, thereby allowing the T-cell to bind to a tumor cell and kill it (Figure 1). These first CAR-T cells did not last long in the body; however, in 1998, it was demonstrated that the introduction of one of the co-proteins necessary for T-cell activation and stimulation (CD28) into the chimeric T-cell receptor could significantly improve the persistence of chimeric T-cells in the body, marking a significant advancement (2nd generation).

In 2002, the first effective CAR-T cells were developed targeting CD19, a marker expressed on B cells. In 2013, the first clinical study was performed in leukemia patients and the first CAR-T was approved in 2017 (Kymriah, tisagenlecleucel) for treating patients with B-cell acute lymphoblastic leukemia (ALL). As of 2022, six CAR-T cell therapies have been approved by the FDA, all of which are active against blood cancers, including leukemia. Despite its hype, CAR-T therapy has been criticized for its severe adverse events, ineffectiveness towards solid tumors, and high cost (upwards of $450,000). There are also concerns about whether CAR-T therapy may cause T-cell cancers, such as lymphoma.

Cell therapies are reviewed and approved by CBER within the FDA under the PHSA. Products submitted to CBER submit a Biologics License Application (BLA), whereas products submitted to CBER submit a New Drug Application (NDA). Regulation of cell therapy has largely remained unchanged, with the exception being approval of Regenerative Medicine Advanced Therapy designation within the 21st Century Cures Act, which created an expedited program for innovative regenerative products, similar to breakthrough designation, which includes *"cell therapies, therapeutic tissue-engineering products, human cell and tissue products, and combination products using any such therapies or products; and may include gene therapies and genetically modified cells, including genetically modified somatic cells,"* as a way to bring experimental cell therapies to patients faster. More on the regulatory aspects of cell therapies is discussed in the chapter on Regulatory Considerations.

Gene Therapy

Gene therapy is the administration of genetic material to treat or prevent disease by modifying a person's genes. Essentially, the genetic material is the drug, and rather than treating the symptoms, as so many drugs do, gene therapy aims to treat the root cause. Gene therapy works by one of the following mechanisms:

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Are mRNA Vaccines Gene Therapy?

mRNA vaccines changed the world – some claim for good, some claim for bad. Their mechanism of action is elegant. In biology, the process for producing proteins, such as enzymes, antibodies, and hormones, is gene → mRNA → protein: genes are transcribed into mRNA, which is then translated into proteins. mRNA vaccines against the SARS-CoV-2 virus work by injecting mRNA coated in a lipid particle directly into the body, which is then distributed to the cytoplasm, where it stimulates the production of an otherwise harmless protein found on the virus's outer shell, thereby bypassing the gene and ultimately stimulating the immune system to destroy the virus.

There has been a lot of discussion and excitement, some of it quite angry, about whether mRNA vaccines, like the COVID-19 vaccines, are gene therapy, implying that these drugs somehow alter a person's DNA or that the vaccine itself is DNA. Some celebrities, like Joe Rogan, have stated that mRNA vaccines are "really gene therapy" [17]. Even scientists have stated that mRNA vaccines are gene therapy [18]. One survey suggests that 1 in 4 people believe that mRNA vaccines change your DNA [19]. But are mRNA vaccines really gene therapy?

The FDA [20] states that:

> *"Human gene therapy seeks to modify or manipulate the expression of a gene or to alter the biological properties of living cells for therapeutic use.*
>
> *Gene therapy is a technique that modifies a person's genes to treat or cure disease. Gene therapies can work by several mechanisms:*
>
> - *Replacing a disease-causing gene with a healthy copy of the gene*
> - *Inactivating a disease-causing gene that is not functioning properly*
> - *Introducing a new or modified gene into the body to help treat a disease"*

Scholarly legal papers on the topic define gene therapy as: *"The intentional modification of a specific somatic gene in order to correct defective genes or to fix an allele's malign function, or the induction of expression of genetic material that compensates for lost or aberrant cellular function, or the introduction of a novel function that would not otherwise be present in the target cell type"* {Sherkow, 2018 #853}. The European Union has specifically stated: "*mRNA COVID-19 vaccines are not considered gene therapies. They are classified and regulated as biotechnology medicinal products."*

That mRNA vaccines are not gene therapy by these definitions seems pretty obvious. mRNA vaccines don't even distribute into the nucleus – they work in the cytoplasm of cells. Still, the impact of someone like Joe Rogan on public opinion, whose podcast averages 11 million listeners, cannot be discounted. Drug developers need to be aware of these definitions when dealing with nonscientists, particularly family members, who ask them questions on this because the possibilities for mRNA drugs go far beyond vaccines and could be a significant class of drugs in the future.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

- **Gene Augmentation:** A cell with a non-functioning gene is administered a functional gene, and the cell becomes functional. This is the most common form of gene therapy and is most effective for loss-of-function diseases, where the body is not producing enough functional protein. An example of this is voretigene neparvec (Luxturna), a gene augmentation therapy for RPE65 in patients with inherited retinal blindness. The RPE65 gene restores production of the RPE65 enzyme, enabling retinal cells to process light normally.
- **Gene Silencing:** A cell with a deleterious gene is administered a blocking gene, silencing the genetic activity of the cell. Drug classes that silence genes include antisense oligonucleotides, interference RNA, and epigenetic silencing.
- **Gene Editing:** A cell with a malfunctioning gene is administered the correct gene, which replaces the malfunctioning gene in the DNA. An example of this is CRISPR technology.
- **Suicide gene therapy:** A gene is administered, and the targeted cell, usually a cancer cell, internalizes the gene and transports it to the nucleus. Once inside the nucleus, the gene encodes for an enzyme that can metabolize a non-toxic prodrug into an active, toxic drug. When the nontoxic prodrug is administered to the patient, it distributes into cells. In cells that contain the enzyme, it converts the prodrug to its toxic form. As a result, those cells will die.

Naked DNA is not administered directly because enzymes in the blood quickly metabolize it, it is poorly taken up inside cells, and it has inherent safety concerns about its use. Today, genetic material is packaged inside a capsid, the protein shell of a virus, which serves as a delivery vehicle to transport it to cells, where it can diffuse into the nucleus. The capsid protects the genetic material from degradation and can carry it directly to cells. Different capsids have varying affinities for different tissues, a phenomenon known as tropism. Most capsids today are viral in nature, raising concerns about their immunogenicity, and all have demonstrated safety issues.

The first gene therapy trial was conducted in 1990, treating two patients with adenosine deaminase (ADA) deficiency, a condition that causes severe immunodeficiency. More patients were treated for various diseases until 1999, when a patient died under controversial circumstances. The field stalled, but was not stopped, for a long time after that. In 2012, the first gene therapy was approved. Glybera (alipogene tiparvovec) received market authorization in the European Union for the treatment of lipoprotein lipase deficiency. It was ultimately withdrawn from the market in 2017 due to concerns over efficacy and low demand. As of 2025, fewer than 2 dozen gene therapies are approved in the US, but worldwide, approximately 700 gene therapy programs are in clinical development. While these numbers seem promising, another cloud has been cast on gene therapy in recent years. Bluebird Bio, a leader in gene therapy, almost went bankrupt in 2025 before being acquired by a venture capital firm. Some gene therapy companies, such as Omega Therapeutics, have gone out of business. Some major companies, such as Pfizer, Takeda, Vertex, and Biogen, have largely abandoned gene therapy as a therapeutic area, at least in part, due to high development costs, safety concerns about the delivery vehicle, and low returns on investment. There were also some notable deaths following administration of gene therapies, leading to questions again about their safety. Gene therapy has been on the ropes before and has come back. Since a typical gene therapy treatment can cost millions of dollars, it remains to be seen how successful gene therapy will ultimately prove to be.

Like cellular therapy, gene therapies are reviewed and approved by CBER within the FDA under the PHSA. Gene therapy differs from other drug classes in several notable ways. Due to concerns about changes to the DNA of cells, particularly sperm and eggs, also known as germline editing, and the possibility of making genetic changes heritable, drug companies must demonstrate that the genetic material in gene therapy does not incorporate into germ cells. Furthermore, patients must be closely monitored for an extended period, sometimes up to 15 years, to ensure the therapy's long-term safety. Lastly, because of early clinical trial deaths involving gene therapy, informed consent with gene therapy is more comprehensive.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 4 Differences Between Brand-Name Drugs and Generic Drugs	
Brand-name Drugs	**Generic Drugs**
Made by the originator company	Not usually made by the originator company
May be marketed under the brand name	Cannot be marketed under the brand name
Exclusivity during patent period	Only available after the patent expires
Active ingredient + excipients	Must have the same active ingredient but may have different excipients
Made using strict quality standards	Also made using strict quality standards
Originator fronts all the major development costs	

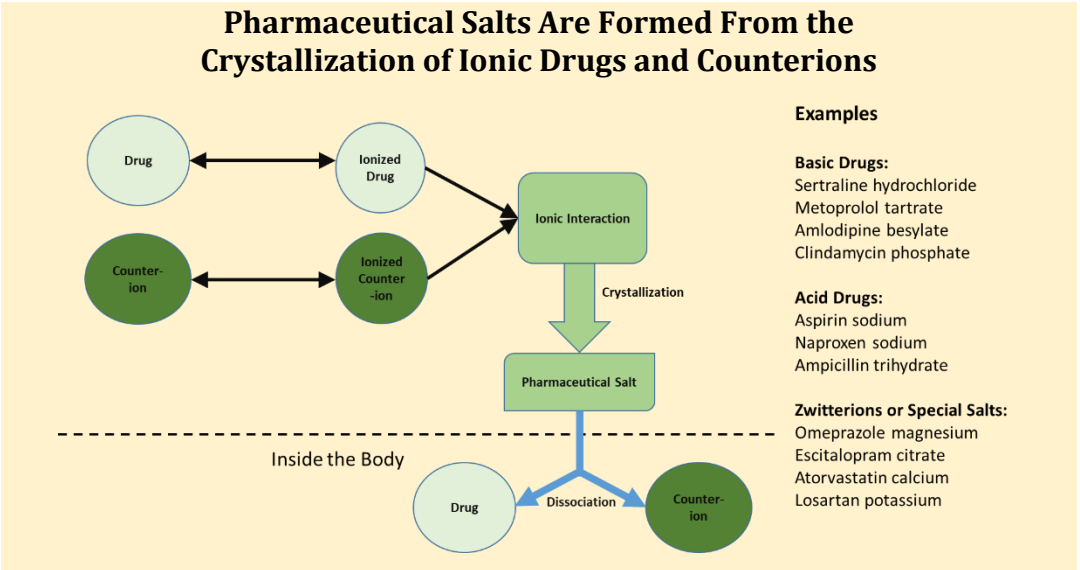
Brand-name Drugs vs. Generics

In most countries, drugs are patented for 20 years. However, with some extensions and clinical trials in specific populations, like pediatrics or rare diseases, patents may be extended for a few more years. Once a drug comes off patent, any manufacturer may begin manufacturing and marketing it as their own product, provided they meet specific regulatory requirements. These drugs are known as generic drugs, or simply generics, and there are similarities and differences between brand-name drugs and generic drugs. Before a regulatory agency can approve a generic drug, it must be chemically identical to the original in terms of active agent, dose, route of administration, quality, performance, and intended use. For example, if a brand-name drug was marketed with strengths of 50 and 100 mg, the generic manufacturer could not make a strength of, say, 75 mg, or if the drug was approved for use in patients with breast cancer, the generic manufacturer could not market it to treat thyroid cancer, or if the brand-name was a tablet, the generic manufacturer could not market a suspension. What can differ regarding a generic drug are the inactive ingredients used in the formulation, known as excipients.

Once approved by a regulatory agency, the product can be marketed under its non-proprietary name, or the manufacturer can market it under its own brand-name. The use of generic drugs rapidly grew after Congress approved the Hatch-Waxman Act in 1984, more formally known as the Drug Price Competition and Patent Term Restoration Act. Before 1984, there was no straightforward regulatory pathway for approving a generic. Generic manufacturers had to conduct the same expensive and time-consuming clinical trials as brand-name manufacturers. Additionally, in conducting these clinical trials, generic manufacturers exposed themselves to legal liability for potential patent infringement. The Hatch-Waxman Act changed all this by creating an expedited pathway for generic approval and certain patent exclusivities for generics upon their approval. It also adjusted patent life for brand-name drugs due to delays in regulatory review and established a patent litigation process for potential patent infringements by generic manufacturers. Generic drugs account for 90% of the 6.5 billion prescriptions written in the US [21].

From a clinical pharmacology perspective, generic drugs must demonstrate chemical and bioequivalence. Chemical equivalence is easy – if the brand-name drug has a tablet with 75 mg of active ingredient, so should the generic drug. A little trickier is whether the generic drug must be formulated with the same salt as the brand-name drug. Some drugs are ionic when dissolved, meaning they are positively or negatively charged. The ionic state of a drug can significantly impact its stability, solubility, dissolution from a tablet to a solution, and pharmacokinetic profile. By

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



formulating the drug as a salt, its chemical stability is often improved, as are its solubility, dissolution, and absorption into the body. For example, Luvox is fluvoxamine maleate (maleate is the counterion for fluvoxamine). Hydrocortisone is often formulated with an acetate counterion, i.e., hydrocortisone acetate. Once in solution, salts rapidly dissociate into their components, and the counterion plays no role in the pharmacodynamic effect of a drug. Generic manufacturers may alter the counterion, but regulatory agencies may require additional tests and studies to ensure the safety, effectiveness, and equivalence of the generic drug.

Bioequivalence, discussed in much more detail in the chapter Principles of Pharmacokinetics, is defined as follows: two formulations are bioequivalent if they have the same rate and extent of absorption at the site of action [22]. This assessment is made by comparing the concentration-time profiles of different formulations of the drug in a biological matrix, such as plasma or serum. If two formulations are bioequivalent, they behave pharmacologically the same (Figure 2). Under certain circumstances, companies may request a biowaiver from regulatory agencies, which allows the generic manufacturer to substitute in vitro tests for clinical studies in humans.

Although the Hatch-Waxman Act has been amended many times, a significant milestone for generics was the Patient Protection and Affordable Care Act, more informally known as Obamacare, approved by Congress in 2010, which allowed generic versions of biological drugs and granted 12 years of exclusivity to brand-name manufacturers before a generic version of the drug could be approved. The difficulty with allowing a generic version of biologics by the FDA was that before the Act, there was no formal regulatory pathway for approval – the generic pathway, as defined by law, was for small molecules.

Generic manufacturers had to show chemical equivalency, but for biologics, that was impossible. Two biologics made by different cell lines are never exactly the same. Hence, generic biologics are referred to as biosimilars. They are similar, but not the same as, brand-name drugs, and they must demonstrate no clinically meaningful performance differences from the brand-name drug to be approved. The EU was ahead of the US in establishing a regulatory framework for approving biosimilars. By 2015, the EU had more than four dozen approved biosimilars, while the US had just approved its first one. The first FDA-approved biosimilar was Sandoz's version of filgrastim, initially made by Amgen and marketed as Neupogen. To denote that Sandoz's version was not the same as Amgen's version, Sandoz's version was given the non-proprietary name filgrastim-sndz, while Amgen's version remained simply filgrastim. Sandoz then launched the product as Zarxio. The approval of biosimilars has skyrocketed, with the FDA approving its 50th biosimilar in 2024.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

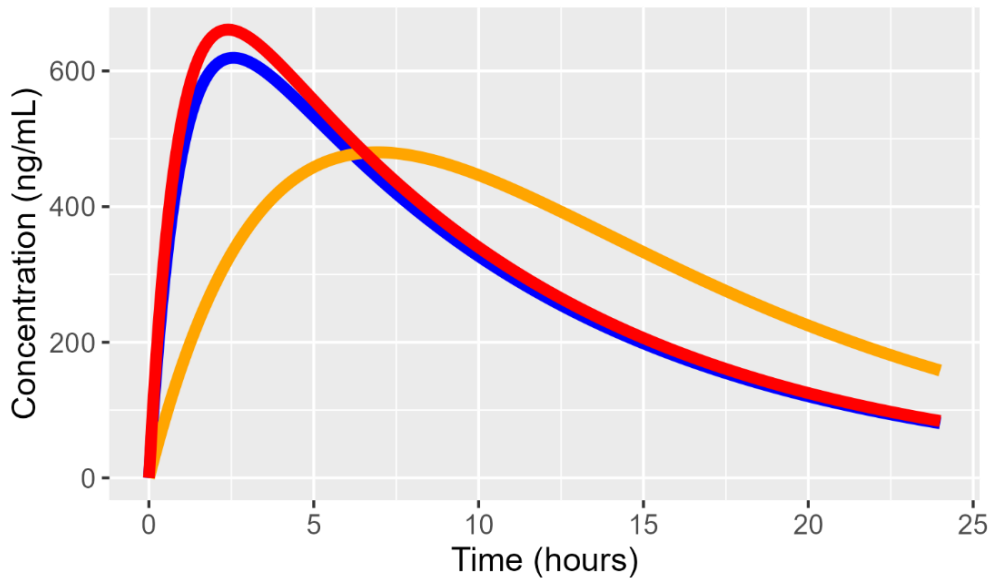
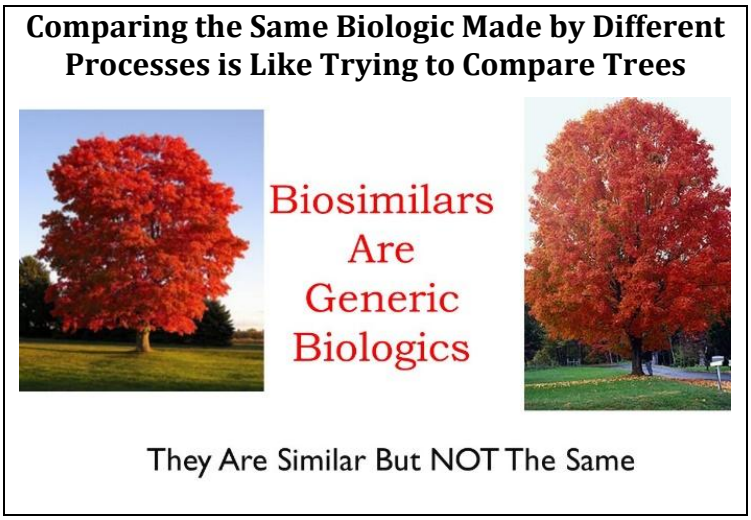


Figure 2. Concentration-time profiles for three formulations of the same drug in the body. The blue profile is the brand-name formulation, while the red and orange profiles are generics. The red profile is bioequivalent to the brand-name profiles, but the orange profile is not because it has a different rate and extent of absorption compared to the other two profiles.

The Hatch-Waxman Act was approved because it was believed that generic drugs would decrease consumer drug costs and that, as more generic drugs became available, drugs would become even cheaper. The same was true for the Affordable Care Act and the biosimilar approval pathway. A natural question then is, did Hatch-Waxman work? Do generic drugs result in decreased costs for consumers? It has been estimated that generic drugs saved Americans $373 billion in 2021. To put this number in perspective, it was estimated that consumers spent $324 billion on prescription drugs in the US in 2015, of which $7 billion was spent on biosimilars. That's the good news.

On the downside, there have been increased reports of poor pharmaceutical quality from overseas generic drug manufacturers, which can lead to potential therapeutic failures. Since the FDA has insufficient inspectors to do quality checks at overseas manufacturing sites, this could lead to manufacturing issues with generic drugs. For example, in 2015, the FDA sent a warning letter to Dr. Reddy's Laboratories, one of the largest generic drug manufacturers, requiring the recall of certain lots of the antiepileptic drug divalproex, the cholesterol-lowering drug atorvastatin, and the antihypertensive drug amlodipine [23]. Furthermore, because the



Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

profit margins for generic drugs are low, some brand-name manufacturers have left the generic business to focus on brand-name pharmaceuticals, thereby reducing the number of generic manufacturers for certain products. In some cases, this has reduced the number of generic manufacturers to a single manufacturer for certain drugs, with unintended consequences. In 2015, consumer outrage fueled numerous news stories when Valeant Pharmaceuticals increased the price of Nitropress (nitroprusside), a decades-old medication used to lower blood pressure, from $215 per vial to $881. Another Valeant drug, Isuprel (isoproterenol), increased from $180 to $1,472 a vial. Valeant bought both drugs, which gave them a monopolistic position, and then increased their price many times higher than the price before they were acquired [24]. This led to patient anger, death threats against the CEO, bad press, and having to testify before Congress to justify their price increases. Ultimately, Valeant's CEO was fired.

Paracelsus (1493-1541), The Father of Toxicology

"dosis sola facit venenum – the dose makes the poison

Figure courtesy of Wikipedia.

The Dose Makes the Poison

Paracelsus (1493-1541) was a German-Swiss scientist and is considered the father of toxicology, the branch of pharmacology concerned with the adverse effects of chemicals on the body. He originally stated, *"All things are poison, and nothing is without poison; only the dose makes a thing, not a poison,"* which is often simplified to "The dose makes the poison." This simple statement is the founding principle of toxicology. It means that everything is a poison; it's just a matter of dose. And when we say everything, we mean everything, including air and water.

You may have noticed that the adverse event section in package inserts provided when you receive a new prescription is lengthy. These potential side effects can sometimes appear to fill a small telephone book. It's been argued that these lists are long to minimize potential legal liabilities. We have all seen the commercials from lawyers advertising to you to call them if you have had a side effect after taking a drug, like if you have suffered the loss of your jawbone while taking a bisphosphonate for the treatment of osteoporosis. Class action lawsuits are meant to recover damages against drug companies for these severe adverse events. Listing these adverse events on

The Dose Makes the Poison- Water Intoxication

Daily Mail

High school football player, 17, dies from water intoxication after drinking FOUR GALLONS of fluids to stop cramps during practice

- Zyrees Oliver's family switched off his life support on Monday
- The teenager collapsed at home in Georgia last week after drinking two gallons of water and two gallons of Gatorade to stop cramps during practice
- Doctors said he suffered brain swelling and was left brain dead
- The family of the teenager, who was a top student and promising athlete, are now raising money to fly his body back to New Jersey, where he is from

By LYDIA WARREN

PUBLISHED: 11:32 EST, 12 August 2014 | UPDATED: 11:59 EST, 12 August 2014

Every summer, high school students die from drinking too much water, often after extreme workouts. It is often not recognized that drinking more than 4 liters of water daily is dangerous, and that consuming 6 liters can be fatal. Excessive water intake can disrupt the body's electrolyte balance and negatively impact brain function, potentially leading to cerebral edema.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

the package insert can limit the manufacturer's liability because it can then be argued that the consumer was warned of the risk. As a society, we have lost touch with Paracelsus's adage. We need to remember that all drugs are potentially dangerous, even ones that are over the counter.

Summary

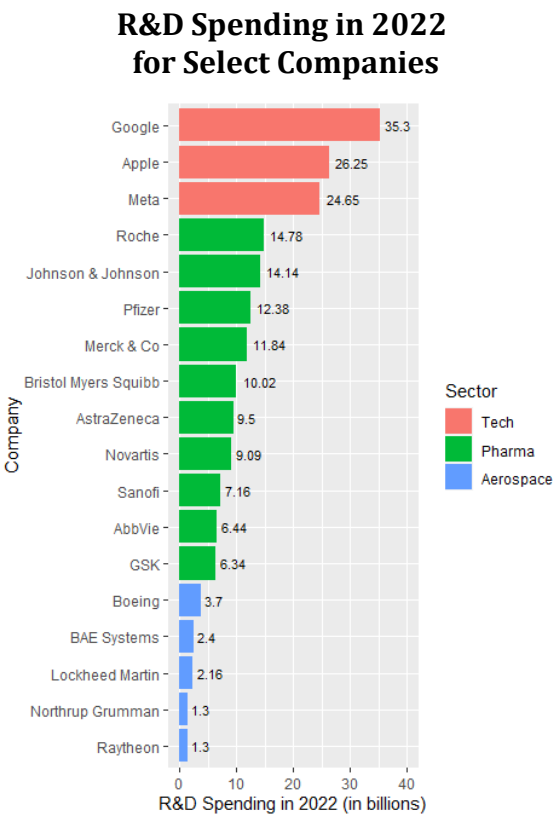
Drugs are an integral part of our lives. It's hard to imagine not having them when we are sick. As patients become more educated about medicine and the options available to them, it is equally imperative that they also become more informed about the drugs they are taking and other medications that may be suitable alternatives. Only then can they make an informed decision. Since 1968, pharmaceutical manufacturers have provided consumers and healthcare professionals with information about their drugs through package inserts. This helps physicians make informed dosing decisions and provides consumers with information about the medication's risks and benefits. Today, these inserts are lengthy documents. The FDA has attempted over the years to simplify them and make them more accessible to consumers, but they are still tedious to read. They are written primarily for doctors and scientists rather than the general public. Hence, the burden is on the public to become more educated if they hope to understand them. Now that we know what a drug is, the remaining chapters will hopefully help decipher the contents of the inserts so that consumers may understand what is written in them.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 2

A Brief History of Drug Discovery

Before the rise of the tech industry in the 21st century, the pharmaceutical industry was the single largest funder of research and development (R&D) in the US Even today, as a percentage of revenue, pharmaceutical companies spend more on R&D than any other sector in the US, including the tech industry, with pharmaceutical companies reinvesting ~25% of their revenues in R&D. For the 10 years between 2010 and 2020, Roche, Pfizer, and Johnson & Johnson each spent more than $80 billion for R&D [27].



Data redrawn from [25, 26].

Now, ask yourself: how much does it cost to develop a single drug? Is it less than $500 million, between $500 million and $1 billion, or more than $1 billion? That is a trick question. Depending on how economists calculate a pharmaceutical company's costs, labor, and expenses, the cost can range from $161 million to $4.5 billion (as of 2019) [28]. That's a billion with a 'b.' Regarding the cost per drug per R&D expenditure, Novartis had 12 drug approvals between 2016 and 2021. During that time, they spent $44 billion on R&D expenditures, which naively translates to $3.6 billion per drug. This is the equivalent of the salary of the top 200 players (out of 484) in the National Basketball Association for 2023 or 600 Super Bowl commercials!

One last question: how long does it take to get a new drug to market, from discovery to approval? The time from the start of studies in humans to approval averages about 9 years [29]. In total, including all animal studies conducted before the initiation of human studies, the time from drug discovery to approval is approximately 12 to 15 years, assuming the drug even reaches that stage. The odds are strongly against any newly discovered drug ever reaching the market. For compounds that reach animal testing, only about 10-20% ultimately make it to market [30]. For some diseases,

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

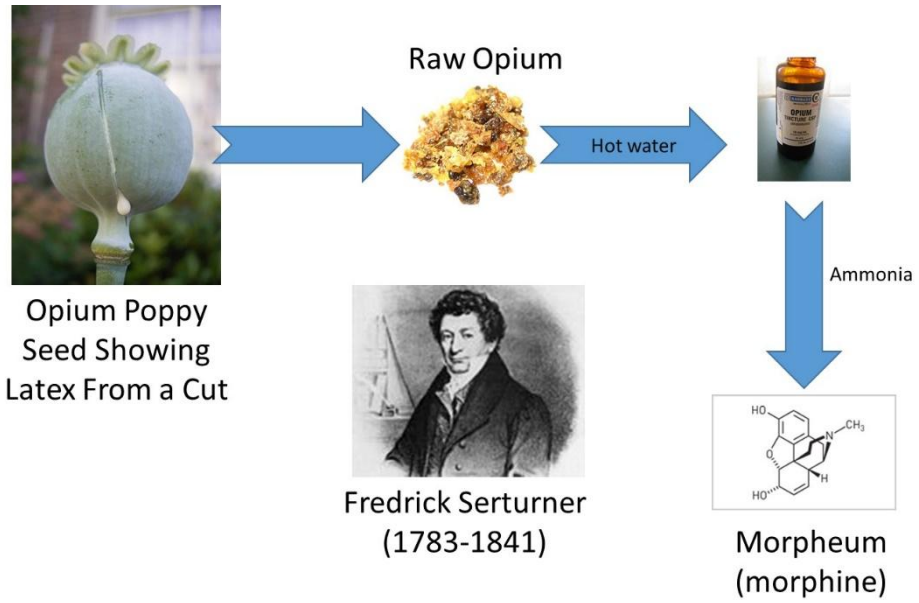


Figure 3. Opium is the dried latex from the opium poppy plant and contains about 12% morphine. Fredrick Serturner, shown in the middle, was a German chemist who was the first to isolate a medicinal product from a plant. In the early 1800s (the exact year this was done varies by source), Serturner purified morphine from raw opium extract by solubilizing it in hot water and precipitating it with ammonia. The extract was 10-times more potent than opium tincture. To demonstrate its activity, Serturner tested the extract on himself and three boys (imagine if this were done today!) with near-disastrous consequences. Subjects experienced 'severe narcosis,' 'pain,' and 'exhaustion.' Based on the dream-like state Serturner experienced, he named the extract 'morpheum' after the Greek god of dreams Morpheus. The name used today, morphine, wasn't coined until 1818 when French chemist Joseph Gay-Lussac translated Serturner's publication into French and recommended that the names of all compounds derived from plant bases end in '-ine.' Hence, he changed the name of morpheum to morphine, which Serturner disagreed with. It wasn't until 1925 that the structure of morphine was identified, and it wasn't until 1952 that morphine was first chemically synthesized.

the success rate is even more dismal, with cardiovascular, psychiatry, and oncology drugs showing success rates of around 5 to 6% [29].

Why is it so hard to bring a drug to market, and why does it cost so much? There are plenty of reasons. One is that the regulatory landscape has undergone significant changes over the last 50 years. Increased regulations lead to longer development times and higher development costs. The number of studies a company must conduct and the number of patients it must study to prove a drug is safe and efficacious, can be quite large and challenging to conduct. The Phase 3 trial conducted in the intended patient population under the proposed dosing regimen is referred to as a confirmatory or pivotal study. In a somewhat dated review of 161 drugs approved in the EU from 2000 to 2010, it was estimated that the median number of patients treated with the drug in the confirmatory trial was 1708, but 16% of medicines had more than 5,000 patients in their confirmatory studies [31]. In cancer patients, it was reported that from 2006 to 2010, 158,810 patients participated in 1335 clinical studies, meaning that, on average, each study enrolled 120 participants [32]. Furthermore, the number of regulations continues to increase. It's challenging to grasp how regulations accumulate over time. Researchers at George Mason University's Mercatus Center and the Regulatory Affairs Professional Society reported that the number of regulatory requirements a regulatory professional needed to know increased from 16,239 in 2000 to 18,777 in 2012, representing a 15% rise. At the

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

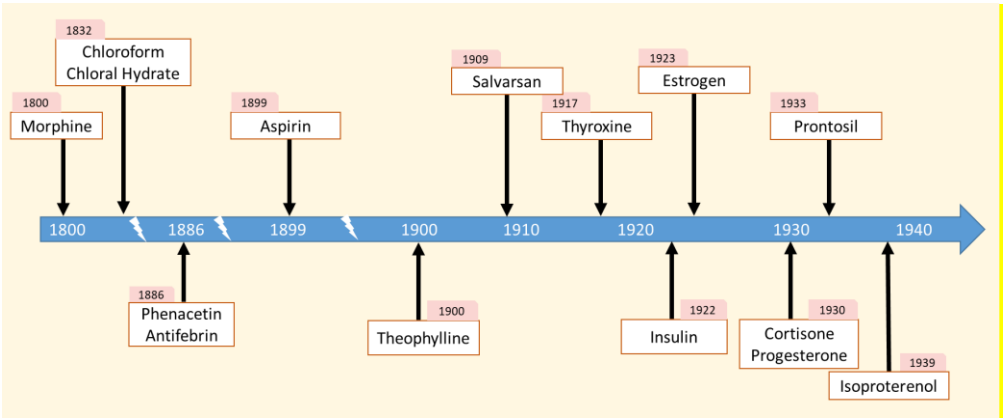


Figure 4. Timeline of important drugs discovered from the 1800s to World War II.

same time, the number of words in the Code of Federal Regulations dealing with the FDA increased by 166,892, which is about equal to the number of words in Charles Dickens's novel "Oliver Twist" (and maybe just as boring).

Another reason is that it is harder to find a drug with suitable properties to be a drug with appropriate safety and efficacy. It's been said that if aspirin were submitted to the FDA as a new drug today, it would be denied approval due to too many side effects. That is true. Every year, the safety and efficacy hurdle increases, making it harder to find suitable drugs. There are other reasons as well, the point being that drug development is long, hard, and expensive. The next time you hear someone complain about the high costs of drugs, maybe you will think about this chapter.

Consider This

Out of every ~ 5,000 new chemically made small molecule compounds identified during discovery, 15-30 are tested in animals, Five are considered safe for testing in human volunteers, and only 1 of these compounds is typically approved as a marketed drug.

People complain all the time that prescription drugs cost too much. Each pill costs only pennies to manufacture, yet pharmaceutical companies charge substantial amounts for each one. That is true. But behind each pill are years of research, studies involving thousands of patients, and thousands of hours of manpower. We wouldn't ask Microsoft to give away the next release of Office for free, but it's the same argument. Microsoft hardly sells any versions of Office on DVD anymore. It's all electronic downloads from the internet. We wouldn't ask them to give Office away for free simply because it costs them only pennies to maintain the servers that deliver Office. Instead, we pay Microsoft the hundreds of dollars it costs to develop the new version.

This chapter provides an overview of the history of drug discovery. Not every drug can be covered in this section, but what is covered represents my attempt to highlight what I consider to be the most important drugs. This section is divided into four parts: drug discovery before World War II, from World War II to 1980, from 1980 to 2000, and from 2000 to the present. Figure 4 presents a timeline of the first section, from the 1800s to World War II.

Before World War II

Until the 1850s, the primary source of drugs was natural products. Plants, fungi, animals, insects, reptiles, and minerals were used to cure disease or relieve symptoms. Pharmacologically active substances derived from natural products include opium tincture from the poppy plant, caffeine from tea and coffee, nicotine from tobacco plants, digitalis and digoxin from the foxglove plant, and

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

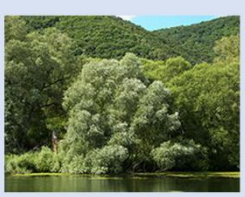
salicylates from the bark of the willow tree. The basic idea was to grind the plant with hot water or alcohol, then drink it like you would tea.

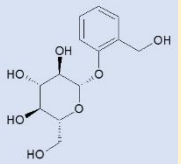
Modern drug research did not fully start until pharmacology became its own scientific field and chemistry had reached a level of maturity that could be applied to outside fields. This started to occur in the early to mid-1800s with the isolation of morphine from the poppy plant by German chemist Joseph Serturner (Figure 3). One of the first pharmacologically active chemicals synthesized by man was the sedative chloral hydrate (sometimes called a Mickey Finn or knockout drops when spiked into alcohol) by German chemist Justus von Liebig in 1832, who also discovered that in alkaline solution, chloral hydrate would decompose into chloroform, which could be used as an anesthetic. With the emergence of analytical chemistry at the end of the 1800s, individual bioactive compounds could be isolated from mixtures with a relatively high degree of purity.

The entire pharmaceutical industry and the chemical synthesis of drugs can be traced back to the textile and synthetic dye industries. During these formative years, and for more than a century after that, luck or serendipity often played a crucial role in identifying new drugs. A rich source of aromatic compounds, like benzene, phenol, and naphthalene, was in the tar remaining after coal distillation. Napthalene had long been used as an internal antiseptic. In the 1880s, naphthalene was studied as a treatment for intestinal worms. After administering naphthalene to a patient, an antipyretic effect was observed, but no effect on the worms was noted. Upon closer examination, it was discovered that the pharmacist had made an error, accidentally dispensing acetanilide instead of naphthalene. When naphthalene was dispensed correctly and used in a different patient, their worms were treated, but their fever was not affected.

The scientists wasted no time publishing their results, which were seized upon by the German chemical manufacturer Kalle and Company, who marketed acetanilide as Antifebrin, the first brand-name drug prescribed by physicians [33]. At the same time, Bayer was heavily invested in dye manufacturing and, seeing the kind of money Kalle and Company were making, faced with tons of para-nitrophenol, a waste product from their manufacturing processes, the Head of Research at Bayer challenged his chemists to turn it into a drug. The result was acetophenetidine, another antipyretic, which was marketed as Phenacetin. Being safer than Antifebrin, Phenacetin brought in enormous sales for Bayer and was sold for

Salicylate Use to Treat Fever and Pain Dates Back to Ancient Egypt





Since ancient times, the bark of the white willow tree (Salix alba), which contains salicin, has been used to help treat fever and pain. Ludwig van Beethoven's death in 1827 is believed to have been partially due to renal failure from the composer's uncontrolled use of willow tree bark for his aches and pains.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

almost 100 years, turning Bayer into one of the leading pharmaceutical manufacturers.

Modern pharmaceutical companies did not begin until aspirin was discovered [34]. The use of salicylates dates back to ancient times, when willow bark was used to reduce fever and treat pain. While you might think that the discovery of aspirin traces back to this, it does not. Instead, the story is much more convoluted. While it doesn't have the accidental discovery component like APAP, aspirin's discovery was based on cold, hard science and the power of chemistry. In 1794, the first home was lit with coal gas, and by 1801, the entire city of Birmingham, England, was lit by coal gas lamps. While this was great for cities, it had one major downside: they produced vast amounts of worthless coal tar.

We've already seen how coal tar was used to discover acetanilide. Early attempts to distill coal tar into ingredients isolated an acidic compound, phenol, in 1842. Phenol was soon found to be a useful disinfectant by the English surgeon Joseph Lister (Listerine, anyone?), and it is still used today. However, phenol is corrosive over time, and alternatives to phenol were being sought by the 1860s.

One of the leading German chemists of the time, Hermann Kolbe, treated phenol with carbon dioxide in the presence of sodium, forming salicylic acid, which also possessed antiseptic properties and was less corrosive than phenol. Although salicylic acid never replaced phenol, it did find use as an internal antiseptic when it was believed that many diseases were caused by gut pathogens. Then, when Swiss physician Carl Buss administered salicylic acid orally to a typhoid patient, the patient's temperature declined, leading to its use as an antipyretic. Around the same time, it was noticed that salicylic acid had a beneficial effect on patients with rheumatoid arthritis.

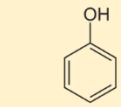
But salicylic acid was hurtful to the stomach, leading to ulcers and profound stomach pain with long-term use. In 1896, Arthur Eichengrun at Bayer (whose job was to discover new products) sought to make a series of phenol esters under the assumption that masking the phenol would protect the stomach and that once in the alkaline conditions of the gut, the ester would dissociate into the free acid. Eichengrun tasked Felix Hoffman with making acetylsalicylic acid (ASA) in 1897. Despite tests showing ASA was superior to any other salicylate, the pharmacologist in charge of evaluating the compounds, Heinrich Dreser, rejected it as a drug candidate, believing ASA would be toxic to the heart.

German chemist Felix Hoffman (1868-1946) has the notable distinction of being credited with creating one of the most useful and dangerous drugs of all time. In 1897, Hoffman invented acetylsalicylic acid, more commonly known as Aspirin, and a short time later, recognizing from his experience with Aspirin that adding acetyl groups to compounds tended to make them less toxic, he acetylated morphine and rediscovered diacetyldihydromorphine, or as it is more commonly known – Heroin. Having got its name from "heroic" or "strong," Heroin was sold as a non-addictive morphine substitute (which was clearly not true) and as a cough suppressant.

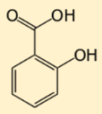
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Who Really Discovered Aspirin?

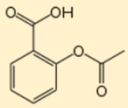
Indeed. Who discovered aspirin is a story unto itself and, to some extent, depends on what you mean by 'discover.' Remember the pharmacologist who rejected ASA as a potential drug, Heinrich Dreser? Bayer management asked him to publish the results of his tests to increase Aspirin's scientific credibility. He did so in 1899, but he omitted the names of Eichengrun and Hoffman as authors from the paper, failing to indicate how ASA was synthesized or developed. Whether he did this purposefully or by accident is not exactly known. To add further insult, ASA was deemed unpatentable under German Patent Law because, under German law, only new processes, not new compounds, are patentable. Hence, Eichengrun and Hoffman would not receive any royalties from the sales of Aspirin under the terms of their contracts, but Dreser set up his contract so that he would receive royalties on sales of any compounds tested in his lab. While Eichengrun and Hoffman would receive nothing, Dreser was about to get rich.



Phenol



Salicylic acid



Acetylsalicylic acid
(Aspirin)

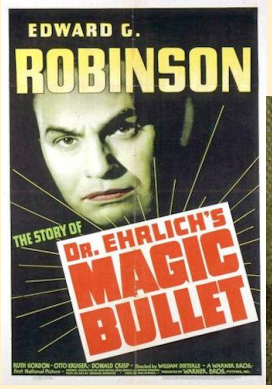
The first account of how ASA was developed did not appear until 1933 in a book on the history of chemical engineering. In a footnote, Felix Hoffman was credited with synthesizing ASA (based on personal communication between Hoffman and the author) and, for decades, claimed to have discovered aspirin because he wanted to help his rheumatic father, who was having difficulty tolerating the large dose of sodium salicylate he needed. This story remained intact for nearly 20 years. But in 1949, after the Nazis lost World War II, on the 50th anniversary of the discovery of ASA, Eichengrun published his account, where he implied that the Nazis rewrote the history of ASA to hide the fact that a Jew had developed ASA. In 2000, historian Walter Sneader did a reappraisal of original archive material. He concluded that the version reported by Eichengrun, the history presented in this book, is the most reliable and that the anecdotal story given by Hoffman was not dependable (at that point, Hoffman was in his mid-60s and his memory might not have been what it used to be) [35]. Bayer disputes Sneader's conclusions. Whether Hoffman discovered ASA independently or under Eichengrun's direction remains a matter of controversy.

Eichengrun then secretly tested ASA on himself and gave it to Berlin physicians for further testing in humans. In addition to showing antipyretic activity, one patient experienced profound pain relief from a toothache, and the analgesic effect of ASA was then confirmed in other patients. Eichengrun then presented the results to Bayer management. They agreed to conduct independent clinical trials to confirm Eichengrun's results. The results convinced Bayer in 1899 to market ASA under the name Aspirin, which was derived from 'a' for 'acetyl,' 'spir' from *Spiraea ulmaria,* which is the botanical name for Meadowsweet, a plant rich in salicin, and 'in', which was a common drug ending at that time. What's interesting about aspirin and APAP is that, given their safety profiles, neither would be approved today by regulatory agencies.

Let's now return to the 20th century. In the first half of the 20th century, advances in organic chemistry enabled scientists to synthesize compounds that did not occur naturally. Pharmaceutical companies began hiring numerous chemists to synthesize compounds, hoping that one of them would

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

prove beneficial. At that time, researchers did not know how drugs worked in the body. They could test them and show that they had an effect, but did not understand why. In the early 1900s, the receptor concept was first proposed, although it wasn't called that at the time, and it took decades for the hypothesis to be universally accepted. The receptor concept is discussed in greater detail in the chapter on Pharmacodynamics.



How many pharmacologists have had movies made after them? Paul Ehrlich proposed that drugs could be designed to target microbes specifically without harming the body. He called these drugs 'magic bullets.' This movie, starring Edward G. Robinson, was made in 1940. Ehrlich's discovery of salvarsan in 1909 for the treatment of syphilis was the first 'magic bullet' and was the basis for chemotherapy [36].

In brief, the idea is that drugs target specific molecules in the body called receptors, and, as such, if you could selectively make a drug that interacts with the receptor, you could develop so-called "magic bullets" to cure the disease. Think of the drug as a key and the receptor as the lock. However, in the first half of the 20th century, the discovery of chemically synthesized drugs was still largely a trial-and-error process.

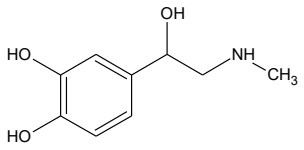
Additionally, during this period, chemists' ability to isolate and purify natural compounds improved. In 1901, epinephrine was isolated and marketed as a vasoconstrictor to relieve bronchoconstriction in asthmatics. The hormone thyroxine was isolated in 1917 and subsequently marketed for the treatment of thyroid deficiency. Insulin was isolated and purified in 1922 from pancreatic glands, and within a few years it was marketed by Eli Lilly & Company worldwide to treat people with diabetes. In 1923, the sex hormone estrogen was purified from pregnant mare urine and then used to treat menopause (this treatment is still used today with Premarin). Ten years later, progesterone was isolated; in 1930, cortisone was isolated.

It was also in the 1930s that the first antibiotics were produced [37]. Using an azo dye called Prontosil that was used to color animal fibers, Gerhard Domagk and his research team found that it could protect mice injected with streptococcus bacteria. In 1935, Prontosil was used to treat 26 women with life-threatening infections. That same year, Daniel Bovet and colleagues reported that Prontosil itself was inactive and only became active after it was metabolized to another compound, sulfanilamide. Compounds that are inactive but require metabolic activation to become active are called prodrugs. Both Domagk and Bovet won the Nobel Prize for their research. From this discovery, a sulfa drug craze hit, and thousands of sulfa drugs were synthesized to find another efficacious drug.

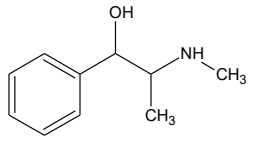
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Sulfa drugs also marked the first drug allergy, where about 2% to 3% of patients taking a sulfa drug develop a skin rash or hives, which could be life-threatening. As noted in the last chapter, sulfanilamide was later responsible for the FFDCA of 1938, which required that drugs be shown to be safe before marketing, when in the year prior, 107 people died from a formulation of sulfanilamide prepared using diethylene glycol, an anti-freeze, as a solvent.

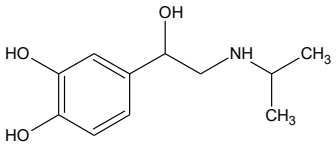
Before there were erectile dysfunction drugs like Viagra and Levitra, Leo Stanley, a (mad?) scientist/surgeon at San Quentin Prison in California from 1913 to 1951, transplanted testicles from executed inmates into aging men in an attempt to "rejuvenate" them [38]. While this may seem unbelievable and completely unethical today, back then, rejuvenation by injecting parts of the body from one person into another or from animals was accepted and something of a fad. Some of these men reported increased sexual potency. This led to the substitution of animal testicles, such as those from boars, goats, and rams. Finally, in 1935, testosterone was isolated, leading to methods to produce it synthetically.



Epinephrine



Ephedrine



Isoprenaline

It was challenging for pharmaceutical companies to patent the compounds they isolated, and they also discovered that many of the compounds extracted from natural sources were not suitable for use as drugs, as they could not be administered orally and their effects did not last long. Chemists then began chemically modifying the natural product to produce more stable compounds with improved efficacy. One of the first such compounds was ephedrine, which has a similar structure to epinephrine. In contrast to epinephrine, ephedrine can be taken orally, and although its effects last several hours, it is less potent. Then, in 1939, C.H. Boehringer of Ingelheim (later to become Boehringer Ingelheim) synthesized isoproterenol (also known as isoprenaline), which had a structure similar to that of epinephrine and ephedrine. Isoproterenol was a significant improvement over both epinephrine and ephedrine and remained the drug of choice for asthma attacks until the 1970s with the approval of salbutamol.

A Tiny History of Insulin

Diabetes affects millions worldwide, and without regular injections of exogenous insulin, they would die early. In 1899, German scientists found that animals developed diabetes shortly after the removal of the pancreas. Later researchers identified that Langerhans cells produced insulin. In 1910, Sir Edward Albert Sharpey-Shafer proposed that a chemical was missing from the cells of diabetic patients, which he coined "insulin," meaning "islet" in Latin. In 1921, a surgeon named Frederick Banting extracted insulin from Langerhans cells and, with his "thick brown muck," as he called it, kept a severely diabetic dog alive for 70 days. He might have been able to keep the dog alive longer, but he had run out of muck. By 1922, researchers had successfully extracted and refined insulin from cattle in significantly larger quantities, and Leonard Thompson, a 14-year-old boy from Toronto, became the first person to receive an insulin injection. These researchers received the Nobel Prize in 1923 for their efforts. The next year, Eli Lilly and Company became the first pharmaceutical company to commercialize large-scale insulin production. Today, patients receive biotechnology-derived insulin, and we no longer use pigs or cattle as a source of insulin.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

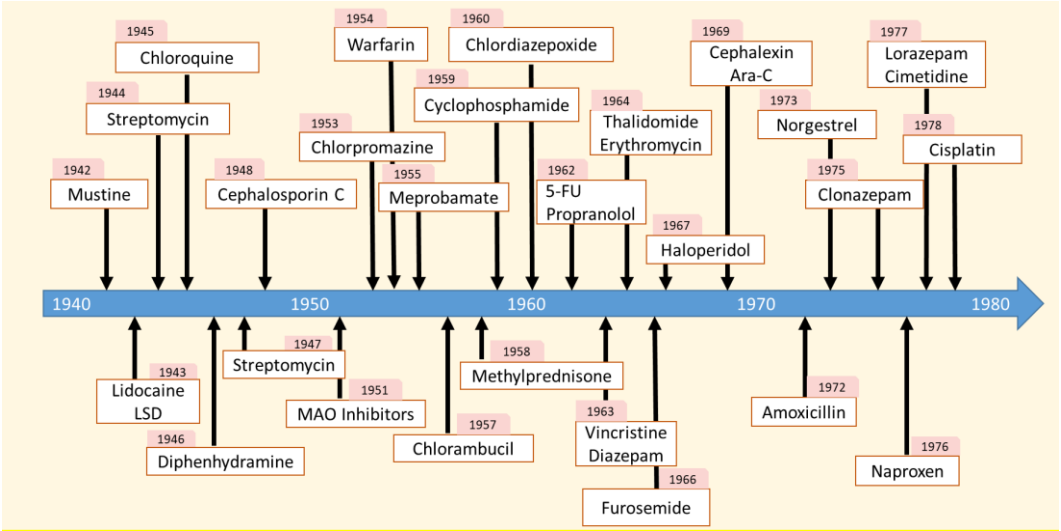


Figure 5. Timeline of important drugs approved or discovered from World War II to 1980.

World War II to 1980 – The Golden Age of Drug Discovery

Figure 5 presents a timeline of important drugs approved from World War II to 1980. With the onset of World War II, drug discovery underwent a significant shift, becoming a central component of the war effort. This development marked the beginning of the pharmaceutical industry as we know it today. Many soldiers were dying from infections they received from battlefield injuries. At the time, sulfa drugs were the only drugs available to treat infections, and they were limited in what they could do. The search was on to find other antibiotics. Although penicillin had been known since 1928, when Alexander Fleming made his amazingly serendipitous discovery, he could not produce it in sufficient quantities to be used medically, and he discarded it [39].

In 1938, Howard Florey of Oxford was interested in the biochemical properties of lysozyme, an antibacterial enzyme originally isolated by Fleming. What happens next is a matter of contention and differs among various sources. In one account, upon discovering Fleming's paper, Florey decided that he would unravel penicillin's antibactericidal action. A few years earlier, Florey had hired biochemist Ernst Chain. In another account, Chain claims he came upon Fleming's paper in the "… oblivion of the literature" and that penicillin and lysozyme were the same or at least similar. This was decades before computers, and looking up prior research was a talent in itself because it required deep diving into old, obscure, and possibly regional journals. Chain claimed he sought to isolate penicillin to prove whether lysozyme and penicillin were identical.

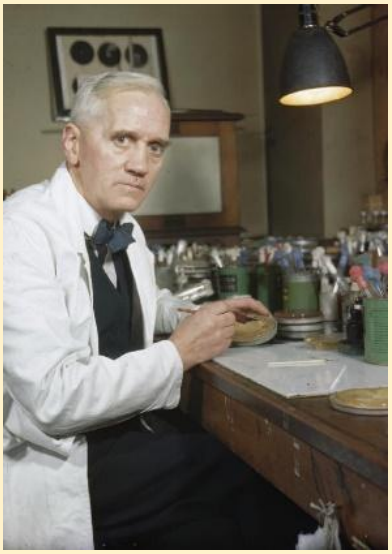
Within a few short months, Britain would declare war on Germany and, with it, all the supplies to support a war, including antibiotics. With a research grant in hand, Florey and Chain sought to isolate Fleming's penicillin. They quickly did so, and biochemist Norman Heatley used every available container he could find to grow as much penicillin as possible and to purify it. To test the activity of what they were isolating, they boldly tested its activity by injecting it in a policeman who had nicked his face in his rose garden and developed a staphylococcus and streptococcus infection that was resistant to sulfa drugs. After five days of injections, the infection began to clear,

Penicillin fungi on an agar plate.

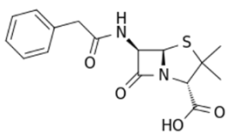
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

but they did not have enough penicillin to cure the infection, and the patient sadly died. Due to the war, British pharmaceutical companies were unable to contribute to the penicillin effort – at this point, no one knew what kind of magic bullet drug they were dealing with.

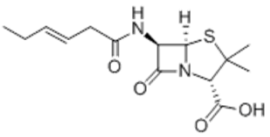
In 1941, before the US entered the war, Florey and Heatley traveled to the US to explore the possibility of collaboration with US pharmaceutical companies in the effort. None would. In the end, Florey and Heatley agreed to collaborate with scientists at the US Department of Agriculture in Peoria, IL, to mass-produce the isolated penicillin. As a first step, the scientists went around town searching for molds to initiate growth. Mary Hunt, a lab assistant, visited local stores in search of moldy fruit. Folks in Peoria started calling her "Moldy Mary." She came upon a cantaloupe that was "a pretty, golden mold" that she knew was "the one!" She labeled it Mold Number 72. Her boss, Andrew Moyer, then placed the mold in a vat of corn steep liquor, and in a second stroke of luck, the mold production increased 200-fold and remained stable. It turns out that the penicillin produced in Peoria, called Penicillin G, was different than that produced by Fleming, called Penicillin F, but they had the same properties. Based on these results, 15 pharmaceutical companies agreed to start mass-producing purified Penicillin G. From January to May 1942, 400 million units of pure Penicillin were produced. By the war's end, pharmaceutical companies were making 650 billion units a month! In 1945, Fleming, Florey, and Chain were awarded the Nobel Prize for their efforts. Heatley was snubbed from that list, possibly because he did not hold a doctorate. To rectify the Nobel Committee's oversight, Oxford awarded its first honorary doctorate to Heatley in 1990.



Alexander Fleming (1881-1955) was a bacteriologist at St. Mary's Hospital in London who is credited with discovering penicillin in 1928, which may be the most serendipitous discovery in medicine. In 1928, after going on holiday with his family, he returned to his lab and noticed that one of the culture plates he had inoculated with staphylococci was overgrown with mold. The staphylococci surrounding the mold were destroyed, but those further away were normal. "That's funny," he said. He later identified the mold as Penicillium. He discovered that the mold could be cultivated and that its antibacterial substance could be extracted. After calling it "mold juice" or "the inhibitor" for almost a year, he coined the substance penicillin. But that is only part of the story. Fleming could never produce penicillin therapeutically because he could not produce it in sufficient quantities. That took the work of Howard Florey, Ernst Chain, and Norman Heatley at Oxford University, more than 20 years later. As Sir Henry Harris of Oxford put it in 1988: "Without Fleming, no Chain or Florey; without Florey, no Heatley; without Heatley, no penicillin."



Penicillin G

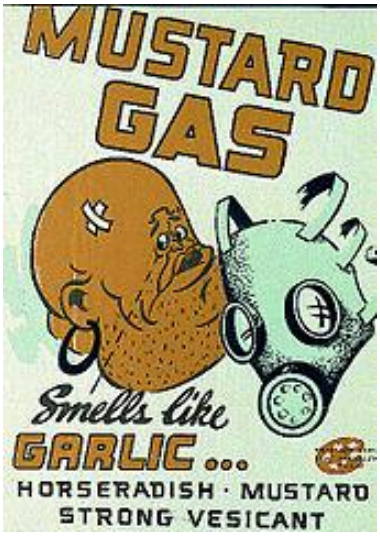


Penicillin F

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

With the success of penicillin, universities and pharmaceutical companies began to look for other antibiotics from natural sources, which came in rapid succession [40]. In 1944, Selman Waksman, who later won the Nobel Prize for his research, isolated the first aminoglycoside antibiotic, streptomycin, and found it was effective in tuberculosis. In 1948, Giuseppe Brotzu isolated the first cephalosporin from a fungus isolated from a sewage outlet along the coast of Sardinia. Also, in 1948, Benajmin Duggar isolated chlortetracycline, also called aureomycin, the first tetracycline antibiotic. Then, in 1952, erythromycin was isolated from *S. erythreus*, a bacterium found in Philippine soil. Other soil antibiotics soon followed: novobiocin in 1955, vancomycin in 1956, and kanamycin in 1957. With chemical modification, scientists at Eli Lilly in 1964 developed this isolate into Keflin (cefalotin), the first cephalosporin drug.

Albert Alexander was the fourth patient to be tested with penicillin, and in 1941, he was the patient used to confirm the activity of the penicillin isolated in Florey's lab.



During World War II, modern-day chemotherapy also began. In World War I, mustard gas killed thousands horribly and was banned afterward by the Geneva Protocol of 1925. However, with the outbreak of World War II, concerns about using these agents increased. Two pharmacologists at Yale, Alfred Gilman and Louis Goodman, who had just written the first edition of what would become the seminal textbook in pharmacology[1], were recruited to study the biological effects of these agents [41]. Based on a literature report of profound myeloid and lymphoid suppression in victims from a German raid in Italy, which resulted in the sinking of the US cargo ship SS John Harvey and the release of the mustard gas bombs it was carrying, Goodman and Gilman thought that these agents might be helpful to treat lymphoma, a blood cancer derived from lymphoid cells. They observed that mustard gas itself was highly unstable; however, when a nitrogen molecule replaced a sulfur molecule, creating nitrogen mustard, it became significantly more stable and could be handled by medical staff. They tested their new compound, which was given the name mustine, in mice and later, in 1942, by injecting nitrogen mustard into a patient with non-Hodgkin's lymphoma [42]. A significant reduction in tumor mass was observed, and although the effect was transient, this was the first study to demonstrate that cancer could be treated pharmacologically.

Shortly after that, based on a study showing that folic acid, a vitamin, increased leukemia cell proliferation when administered to children, Sidney Farber at Harvard in 1948 used folate analogs in children with acute lymphoblastic leukemia and achieved brief remissions [36]. These analogs were aminopterin and methotrexate. Then, in 1951, it was shown that methotrexate has activity in solid tumors in patients with breast cancer. At the same time, George Hitchings and Gertrude Elion at Burroughs Wellcome rationally designed a series of compounds, assuming that synthesizing compounds that resembled the natural building blocks of cells but were not would inhibit cell growth. After determining that cancer cells require purine for DNA synthesis, they developed a series of purine analogs that, when used, would stop DNA synthesis. These drugs, known as antimetabolites, led to

[1] First published in 1941, The Pharmacological Basis of Therapeutics is still the standard text for pharmacology and its 14th edition was published in 2022.

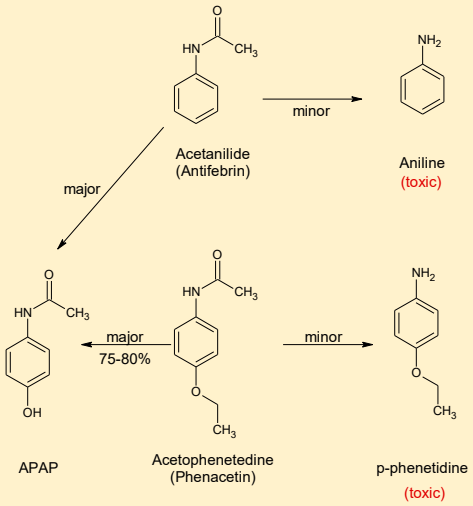
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

the discovery of diaminopurine and thioguanine in 1950. Elion later substituted an oxygen molecule for a sulfur molecule, creating 6-mercaptopurine, which is active against leukemia.

The 1950s also saw the release of one of the most famous drugs of the 20th century – Tylenol (acetaminophen) [43]. Phenacetin and Antifebrin, two older antipyretics from the early 1900s, were not completely safe. Both caused methemoglobinemia in rare cases. Methemoglobin is a form of hemoglobin with a reduced ability to release oxygen into the tissues. High concentrations in red blood cells lead to loss of breath, dizziness, mental status changes, and, in extreme cases, coma and death. Why Phenacetin and Antifebrin caused this severe side effect was studied in 1948 by Julius Axelrod and Bernard Brodie, the former a future Nobel Prize winner and the latter a Lasker Award winner. They found that a small fraction of Antifebrin was metabolized to anilide, a known cause of methemoglobinemia. They also found another metabolite in more significant quantities, N-acetyl-p-aminophenol (APAP). They found that Phenacetin was predominantly metabolized to APAP and, to a minor extent, to p-phenetidine, another metabolite that causes methemoglobinemia.

Although studied as an antipyretic more than 50 years earlier in 1893 by clinical pharmacologist Joseph von Mering, APAP was rejected as a therapeutic because it too caused methemoglobinemia. Such was von Mering's reputation that his findings were not questioned until Axelrod and Brodie showed that APAP had the same antipyretic and analgesic effects as Phenacetin and Antifebrin, but lacked its effect on hemoglobin. They showed that von Mering's findings were wrong. Later studies suggested that von Mering's studies with APAP may have been contaminated with 4-aminophenol, from which von Mering synthesized his APAP, and which also causes methemoglobinemia. McNeil Laboratory used these findings and, in 1955, began marketing APAP as Tylenol Children's Elixir under the generic names paracetamol in Europe and acetaminophen in the US.

Acetanilide (Antifebrin) and Acetophenetidine (Phenacetin) are Both Metabolized to Acetaminophen (APAP)



Tylenol Children's Elixir was Launched in 1955 by McNeil Labs For Children

Marketing directly to physicians and using packaging that catered to children (notice that the box looked like a firetruck; God only knows how many children overdosed on Tylenol thinking they were taking candy; this was before child-safe caps) McNeil Labs sold APAP as a safe and effective alternative to aspirin. This product was soon expanded to other products, and today, it continues to generate sales of over $200 million annually in the US.

The 1950s also saw the introduction of drugs outside of antibacterials and chemotherapy. In particular, the 1950s saw the introduction of drugs affecting the brain. Monoamine oxidase (MAO) is a class of enzymes that oxidize monoamine neurotransmitters, such as norepinephrine and serotonin, two of the brain's major neurotransmitters. In 1951, MAO inhibitors for the treatment of depression were introduced [44]. Starting with the observation that tuberculosis patients prescribed iproniazid became happier, it was discovered that

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

iproniazid had an additional property: it inhibited monoamine oxidase (MAO), thereby inhibiting monoamine metabolism. This observation led to the monoamine theory of depression, which holds that depression is the result of depletion of serotonin, norepinephrine, and/or dopamine in the central nervous system. In theory, drugs that block the depletion or reverse the depletion should ablate the depression. The first class of antidepressants, MAO inhibitors, works by blocking the enzyme responsible for metabolizing these monoamine neurotransmitters, thereby increasing the amount of neurotransmitter in the neurons. However, with these drugs, a new type of severe drug interaction was noted where foods rich in tyramine, like pizza and beer, could lead to a hypertensive crisis because tyramine, which is a neurotransmitter, also increases blood pressure (more on this in later chapters). Inhibiting tyramine metabolism increases its concentration, thereby elevating blood pressure.

Early advertisement for Thorazine from the 1960s.

Drugs for the treatment of schizophrenia were also introduced in the 1950s [45]. After the war, Henry Laborit, a French physician, was interested in the use of antihistamines before surgery to prevent possible histamine release in patients in post-surgery shock. He started with a cocktail of mepyramine and promethazine, both known antihistamines. Over time, he found that the mood of his patients improved and, with promethazine, they required less morphine and were less anxious. Laborit went to Rhone-Poulenc to meet with the chemist who had synthesized promethazine and persuaded him to create a new set of more potent analogs. One of them was chlorpromazine, which affected the peripheral nervous system the least. Physicians found it to be the most effective calming drug to date when tested in patients experiencing shock. Chlorpromazine was released in 1953 under the name Largactil, derived from the terms "large" and "acti" for activity. Following this, Laborit studied the drug in patients with severe burns and psychiatric patients. In one patient, it was shown to have a profound effect on their mania, which led to a study in manic patients. The patients experienced profound changes in mood, thinking, and emotional behavior. Within two years, it was marketed in the US for the treatment of psychosis under the trade name Thorazine because when you take it, it is like being hit by Thor's Hammer. After all, the action is so rapid. The spectacular success of chlorpromazine caused other companies to develop derivatives of it, which eventually led to the family of first-generation antipsychotics: prochlorperazine, fluphenazine, thioridazine, and chlorprothixene.

Along with these discoveries was the introduction of lithium to treat bipolar disorder. Lithium is a natural element on the Periodic Table belonging to the same family as hydrogen, sodium, and potassium. In 1949, Australian psychiatrist John Cade was trying to isolate a compound from the urine of schizophrenia patients that might be responsible for their illness. He needed something to act as a control in his experiments. He believed that since uric acid was helpful in the treatment of gout and was shown to be psychoactive, he used lithium urate as the control. When he administered lithium urate to rats, much to his surprise, they became tranquil, which Cade attributed to the lithium cation rather than the urate anion. Cade then successfully treated manic psychiatric patients with lithium citrate and lithium carbonate, allowing some of them to be discharged from the hospital. Unfortunately, the impact of lithium for the treatment of bipolar disorder took a few years to catch on because of an incident in the US where a salt manufacturer switched from sodium chloride to lithium chloride, and people died as a result of lithium toxicity [46]. Today, however, lithium is used as a "mood stabilizer."

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

While these advances in severe mental illness profoundly changed the practice of psychiatry, nothing compared to the impact that Miltown (meprobamate) had. During World War II, Frank Berger was looking for a preservative for penicillin when he noticed that the compound mephenesin had a tranquilizing effect on rodents. But mephenesin had its own problems. It was short-acting and not very potent. In 1950, Berger and colleagues synthesized an analog of mephenesin that did not have these issues. They called this compound meprobamate. Wallace Laboratories bought the drug and named it Miltown. Launched in 1955, it gained popularity as an anxiolytic.

The anxiolytic Miltown was named after the borough of Miltown, New Jersey, where it was invented.

Miltown was called "Momma's Little Helper" and was advertised in mainstream magazines such as Newsweek, Time, and Reader's Digest, as well as in women's magazines such as Ladies' Home Journal. The Rolling Stones even wrote a song about it, called "Mother's Little Helper," with lyrics like "Doctor please/Some more of these" and "Momma needs something today/To calm her down." Patients flooded doctors' offices demanding the drug. Pharmacies hung signs out their windows saying, "Out of Miltown" or "More Miltown Tomorrow." Reaching such a broad audience made Miltown the first blockbuster psychotropic drug in history, and by 1956, it was the most prescribed medication in America. Consumer Reports estimated that 1 in 20 Americans was taking Miltown, and by 1957, more than 35 million prescriptions had been written, when the total US population was 172 million.

Advertisements like these for Miltown turned it into the first blockbuster psychotropic drug.

This psychotropic revolution and popularization of psychotropic medication continued into the 1960s. Even into the latter half of the 20th century, serendipity continued to play a role in drug discovery. In 1954, while searching for drugs with psychotropic properties, Leo Sternbach, a chemist at Hoffman-LaRoche, went back to a series of heptodiazine analogs he had synthesized in graduate school as synthetic dyes. He synthesized 40 compounds and submitted them for pharmacological activity testing; all were inactive. Chemical analysis revealed that the analogs were not heptodiazine analogs but were benzodiazepine analogs. While cleaning the lab a year and a half later, another chemist found one of the analogs and realized it had not been submitted to the original group for testing. The chemist suggested that it be submitted for testing, and Sternbach agreed. The compound showed activity similar to that of chlorpromazine (Thorazine). This compound was later named chlordiazepoxide and, in 1960, was marketed as Librium, the first benzodiazepine sedative/hypnotic drug. Three years later, Valium (diazepam) was introduced by Hoffman-LaRoche and, from 1969 to 1982, was the top-selling drug in the US. By 1977, benzodiazepines were the most prescribed class of drugs globally.

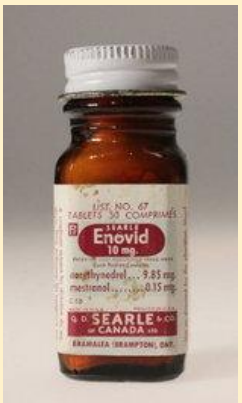
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

These discoveries and new psychotropic drugs led to a revolution and a new golden age of psychiatry, where drugs could treat psychiatric illnesses, transforming mental illness from one of long-term hospitalization to one of community mental health. With Miltown, mental illness shifted from "Blame the Mother to blame the brain," as psychiatrist Eliot Valenstein put it, and dramatically lessened the role of psychotherapy thereafter. What was remarkable about these drugs is that scientists had no idea how these new psychotropic drugs worked. In 1950, brain activity was thought to be entirely electrical, and there was no concept of neurotransmission. This was the age of electroconvulsive therapy and insulin comas. Future Nobel Laureate Arvid Carlsson was mocked by colleagues when he proposed that dopamine might be a neurotransmitter. Developing drugs that affect brain function was a truly fantastic achievement.

Advertisement for Valium from 1965.

The 1960s are remembered for "sex, drugs, and rock n' roll." The drugs part certainly started with the psychotropic revolution in the 1950s, but the sex part started with the introduction of "the pill." In 1951, birth control activist Margaret Sanger was 72 years old and was looking for someone to make for her the "magic pill" that would prevent pregnancy [47]. At a dinner party one night, she met biologist Gregory Pincus, and during their discussion, he surprised her by saying it might be possible to do such a thing with hormones. He agreed to investigate but said he needed funding to do so. Sanger convinced Planned Parenthood to provide the funds for his pilot project and showed that progesterone prevents ovulation in animals. Pincus later worked with Harvard obstetrician and gynecologist John Rock and, later, pharmaceutical company G.D. Searle, to develop an oral form of progesterone. Using a 21-day on/7-day off regimen of synthetic estrogen and progestin, in their efforts to keep contraception as natural as possible, they showed spectacular success in their initial efforts. In their first study of 50 women, none of the women ovulated while on the drug. After some setbacks related to contaminants in the formulation, Searle submitted their product to the FDA for approval, and it was approved in 1957 under the name Enovid.

However, Enovid was approved for the treatment of severe menstrual bleeding, with a warning that Enovid will prevent ovulation. Doctors and women quickly saw what Enovid could do, and within two years, over half a million American women "developed" severe menstrual bleeding and were prescribed Enovid. In 1959, realizing that Enovid could be a huge moneymaker for the company, they re-filed with the FDA, this time as a contraceptive pill, and in 1960, it was approved. Within two years, Searle's monopoly on the market was lost when Syntex released Ortho-Novum in the ever-familiar today DialPak dispenser. Today, there are more than 70 different birth control pills available to women.



Enovid was the first contraceptive pill approved for use in 1960.

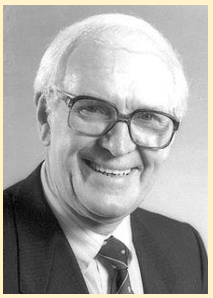
The 1960s also witnessed a profound shift in the treatment of heart disease with the introduction of beta-blockers, calcium channel blockers, and diuretics [48][49]. In the late 1940s, Raymond Ahlquist proposed two types of adrenergic receptors: alpha and beta [50]. When norepinephrine was applied to alpha-receptors, excitatory responses, such as vasoconstriction, occurred, and when it was applied to beta-receptors, inhibitory responses, such as vasodilation, occurred. At the time, this proposal was not universally accepted. In fact, for a while, the original paper proposing this idea was rejected for publication because it contradicted prevailing dogma, and even after publication, it was

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

largely ignored for over a decade. Today, receptor subtypes form the basis of modern pharmacology and drug development.

In 1957, Irwin Slater of Eli Lilly was working with analogs of isoproterenol, a compound discovered in 1939 and known to be a bronchodilator. Using isolated tracheal strips, Slater would bathe the strips in adrenaline to ensure they remained responsive before testing the new analog on them. He noticed that those strips that had been exposed to dichloroisoproterenol (DCI) did not relax when exposed to adrenaline. Neil Moran at Emory University then tested the cardiac effects of DCI and found that it antagonized the increase in heart rate and muscle tension produced by adrenaline. These studies were the first proof that beta-receptors could be blocked by drugs, leading Moran to refer to these drugs as "beta-adrenergic blockers" or simply "beta-blockers."

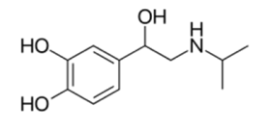
These results intrigued James Black, a Scottish pharmacologist working at Imperial Chemical Industries (ICI) in Great Britain, who, in 1958, was interested in finding drugs that reduced myocardial oxygen demand in hearts compromised by arterial disease. Using Ahlquist's dual receptor theory and recognizing the importance of Moran's work, Black believed it was possible to develop a drug that bound to receptors in the heart, the so-called beta-receptors, without having any intrinsic activity of their own. Black's group then synthesized analogs of DCI, leading to the creation of pronethalol, launched in 1963 as Aderlin, the first non-selective beta-blocker. Pronethalol had the properties Black was looking for, but caused tumors in mice and, as such, was only allowed in humans in life-threatening circumstances.



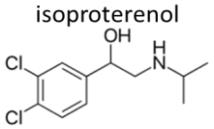
Sir James Black was awarded the 1988 Nobel Prize for discovering the beta-blocker propranolol.

He also discovered cimetidine, the first H2-histamine blocker, which was marketed ten years later for treating peptic ulcers.

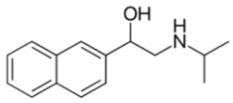
Black and his team then returned to the laboratory and created another series of analogs, which ultimately led to the development of propranolol. Clinical trials showed that it was more effective at lowering blood pressure than pronethalol and had fewer side effects. In 1964, propranolol was marketed as Inderal for treating angina, the first new treatment since the release of nitroglycerin 100 years earlier. It rapidly became the best-selling cardiovascular drug on the market. But Black did not stop there. Faced with the development of sotalol by Mead-Johnson, which was less active than propranolol but had fewer side effects, Black and his team created other propranolol analogs and discovered an even more cardioselective drug, practolol. Practolol was marketed as Eraldin and proved to have its own role in the treatment of cardiac arrhythmias, but it was removed from the market in 1975 for the profound toxicity (blindness, skin rash) it caused. However, differences in their effects led Black to propose the existence of beta-receptor subtypes to explain the pharmacologic differences between propranolol and practolol.



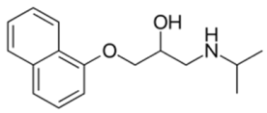
isoproterenol



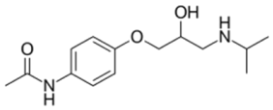
dichloroisoproterenol



pronethalol



propranolol



practolol

The other two drug classes developed in the 1960s that solidified the use of medicinal treatment of heart disease were the discovery of calcium channel blockers [51] and diuretics [49]. Papaverine is an old drug from the 1800s used to treat visceral spasms and vasospasms. Working with a series of papaverine analogs, Ferdinand Dengel of Knoll AG synthesized verapamil, which was found to be a coronary muscle dilator and later found to be antiarrhythmic. Albrecht Fleckenstein then

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

demonstrated that verapamil's inhibition of cardiac contraction could be reversed by adding calcium; he referred to this drug as a "calcium antagonist," a separate and distinct pharmacologic class previously unknown. Another calcium antagonist, nifedipine, was also investigated by Fleckenstein at the same time. With the development of verapamil and nifedipine, the drugs offered new treatments for angina, hypertension, and certain types of arrhythmias.

Modern-day diuretics in the 1960s started with an old friend, sulfanilamide. A side effect of sulfanilamide is its mild diuretic effect; however, it is ineffective as a clinical diuretic because toxicity outweighs the benefit at the doses required to achieve a clinical effect. Starting from sulfanilamide, chemists began building analogs, and from these, many different diuretics were developed and released on the market, including diclofenamide, chlorothiazide, hydrochlorothiazide, and furosemide. The introduction of diuretics offered new treatments for hypertension, heart failure, liver disease, and some types of kidney disease.

The 1960s were also a period of one of the worst disasters in the pharmaceutical industry's history, an event that would forever change the way it conducted business. In 1964, at the start of the anxiolytic craze that was sweeping the world, German pharmaceutical company Chemie Grünenthal released Immunoprin, or as it is more notoriously known – thalidomide [52]. Sold as a sedative/hypnotic and to treat morning sickness, shortly after its release, case reports of deformities and congenital disabilities were reported, and soon more than 5000 cases were reported. The defects were so severe that only 40% of children with them survived. The congenital disabilities seen in Germany were not seen in the US to the same extent because the FDA reviewer, Dr. Frances Kelsey, refused to approve the drug because of the safety concerns she had. Unfortunately, the birth defects seen with thalidomide were not wholly avoided in the US. One of the original provisions of the 1938 FFDCA was that a manufacturer could sell a drug after 60 days if the FDA did not prevent its marketing. About 1200 women were prescribed thalidomide in the US before its complete ban.

At the time, Chemie Grünenthal did nothing wrong and used accepted drug development practices. Teratogenicity testing (whether a drug causes birth defects) was not required or done in the 1960s. As a result of the disaster, the US Congress approved the Kefauver-Harris Amendment to the FFDCA in 1962, which now required drug manufacturers to prove their drug was effective based on adequate and well-controlled clinical studies [53]. The act also had several other provisions to update the original act. These included:

- The act gave the FDA 180 days to approve a drug and required that any drug must have FDA approval before it could be marketed;
- Manufacturers now had to follow good manufacturing processes as set forth by the FDA;
- Transferred to the FDA control of prescription drug advertising;
- Manufacturers had to report any serious adverse events to the FDA after a drug was approved for marketing and
- Controlled the use of generic drugs.

In October 1962, President John F. Kennedy signed the Kefauver-Harris Amendments to the 1938 Food, Drug, and Cosmetic Act into law. Shown 2nd to the left is Frances Kelsey, the FDA reviewer who refused to approve thalidomide in the US and thereby prevented the tragedy seen in Europe.

To this day, the thalidomide tragedy remains; in 2014, eight thalidomide survivors filed a lawsuit against Grunenthal seeking unspecified compensation. Ironically, thalidomide did not go into the dustbin of old, unused drugs. Today, thalidomide is a valuable treatment for erythema nodosum leprosum, a condition that causes red, nodular lesions, fever, weight loss, and general malaise in patients with leprosy. Thalidomide is also used to treat cases of tuberculosis where standard drugs fail. Furthermore, in 2004, Celgene developed an analog of thalidomide, Revlimid (lenalidomide), for the treatment of multiple myeloma. In

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

light of these indications, the FDA has established robust controls and conditions for prescribing to prevent malformations during pregnancy.

Lastly, no discussion of the 1960s is complete without Albert Hoffman and the discovery of LSD (lysergic acid diethylamide) because when you think of "sex, drugs, and rock n' roll," what immediately comes to mind is LSD, which ironically was not invented in the 1960s but decades earlier in the 1930s [54]. Albert Hoffman, not to be confused with the counterculture activist Abbie Hoffman, was a Swiss chemist at Sandoz in Switzerland when, in 1938, he was working with derivatives of lysergic acid, produced by the ergot fungus. Hoping to create a drug that would stimulate the respiratory and circulatory systems, he created 25 analogs of LSD, the last one by mixing LSD with diethylamine, and was named LSD-25. Although the animals administered LSD-25 became highly excited after dosing, Hoffman found LSD-25 to be devoid of pharmacologic activity, and he set it aside for more than five years. But for some reason, Hoffman could not stop thinking about LSD-25, perhaps due to the excited animals, and in 1943, he decided to take a second look at it.

LSD is one of the most potent drugs known to man, taking only micrograms for it to be effective. Hoffman thought he was dealing with a deadly poison when he accidentally ingested LSD-25. When he went home, he had the first accidental "acid trip," which he described as "[being] taken to another world, another place, another time. [His] body seemed to be without sensation, lifeless, strange." In the 1960s, people would buy LSD on "blotters" (example shown above), where they would keep it in their mouth for a while and then swallow.

While re-synthesizing LSD-25, he accidentally ingested a small amount, maybe through his skin or by accidentally touching his hand to his mouth[2], but what happened next lives in memory. He was forced to leave the lab due to "remarkable restlessness, combined with slight dizziness." When he got home, he sat down and, in a dreamlike state, with eyes closed ([he] found the daylight to be unpleasantly glaring), [he] perceived an uninterrupted stream of fantastic pictures, extraordinary shapes with intense, kaleidoscopic play of colors." He initially thought the effects were due to the residual chloroform solvent used in the synthesis, but when he intentionally inhaled chloroform fumes, they did not have the same effect.

Three days later, Hoffman tried again, this time to confirm the results by intentionally ingesting LSD. This day is often referred to as Bicycle Day because he didn't make it home before the psychedelic effects started. Because World War II was ongoing, he had to ride his bicycle home. The LSD effects happened while he was riding, thus making this the first accidental "acid trip." Hoffman continued to research and develop psychedelic drugs, calling LSD "medicine for the soul." In 1947, Sandoz released LSD to the public as a drug called Delsyid for investigational use in psychiatric disorders. LSD use began with psychologists in the 1950s to unblock repressed thoughts. As a result, many psychiatrists and psychologists began using it recreationally. Shortly after its release, the Central Intelligence Agency began experimenting with it, gradually expanding its use in the military and among students. Several prominent counterculture celebrities, like Timothy Leary and Hoffman himself, advocated the general use of LSD for fun. By 1965, Sandoz stopped production of LSD because of complaints and public pressure, and in 1968, possession of LSD was made illegal by the Federal government. This did not last long because, in the 2000s, doctors started using LSD derivatives to treat post-traumatic stress disorder in veterans. In 2022, Oregon became the first

[2] My first job in the pharmaceutical industry was at Eli Lilly and Co. I was told this story, and I don't know if it is true, but it seems true, about a chemist who worked there while they were developing LSD analogs in the 1960s. This chemist had a habit of cleaning his spatula, a small spoon used to weigh out drugs on a balance used to weigh out drugs, by wiping it over his tongue and then drying it on his lab coat. For one particular compound, it apparently caused a bad hallucinogenic episode, an ambulance had to be called, and he was restrained for hours until he recovered. I assume I was told this story so that I would not use his bad habit and clean my spatula with my tongue.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

state to approve psilocybin (an analog of LSD) for therapeutic use.

The 1970s can be marked by two significant events: the War on Cancer and the emergence of biologic drugs. Even though mustard gases were released in the 1940s, the 1950s were marked by a period of pessimism regarding cancer therapies. In the 1960s, oncology as a medical specialty did not exist (it was not established until 1973), and medical universities discouraged research into new cancer drugs. But with the hard-fought efforts of luminaries like Vince DeVita, Paul Calabresi, Gordon Zubrod, Emil Frei, and Howard Skipper, by the end of the 1960s, it was believed that cancer might be curable. In 1971, President Richard Nixon signed the National Cancer Act, which was sold to the American public as the "War on Cancer." With that signing, vast amounts of money became available for cancer research and the study of new cancer drugs.

The day of the first LSD trip, 19 April 1943, is referred to as Bicycle Day because after ingesting LSD, the effects Albert Hoffman experienced on the ride home on his bicycle convinced him he had created a potent psychedelic drug. This LSD blotter commemorates that experience.

However, in 1956, Gordon Zubrod was appointed head of the Division of Cancer Treatment at the National Cancer Institute (NCI). He quickly set up the largest collection of American plants to test for anticancer activity. Sacks of needles, twigs, bark, seeds, and leaves were collected. From these extracts, samples were prepared and tested for their ability to kill cancer cells in vitro. Further work was done to isolate the active component if something proved active. Over 30,000 species were collected. One of these was from the Pacific Yew tree, an old-growth tree in the Pacific Northwest. When tested in 1964, it was found that an extract of the bark was active, but the leaves and twigs were not. After isolating the active ingredient, it was named Taxol [55]. By 1969, 10 kg of pure Taxol had been isolated from 1,200 kg of Pacific Yew bark, but it failed to progress beyond this stage for reasons not reported. If I had to guess, I would say that the compound did not progress further because of the difficulty and low yields in extracting pure Taxol from Pacific Yew tree bark. Even if the compound showed promise, it would be untenable to support human development due to the amount of bark required.

By 1971, the chemical structure of Taxol was identified. Still, the structure was so complex to synthesize that studies continued to use Taxol derived from the Pacific Yew tree bark. By 1980, an order for 20,000 pounds of bark was placed for botanists to collect. Following animal studies, clinical trials in cancer patients were initiated in 1984. The results were promising, and further studies were initiated in patients with melanoma and ovarian cancer. This time, 30% of patients with ovarian cancer responded, a remarkable achievement. Based on the current incidence of breast cancer in the population, it was estimated that 360,000 Pacific Yew trees would be needed to supply the drug *per year*!

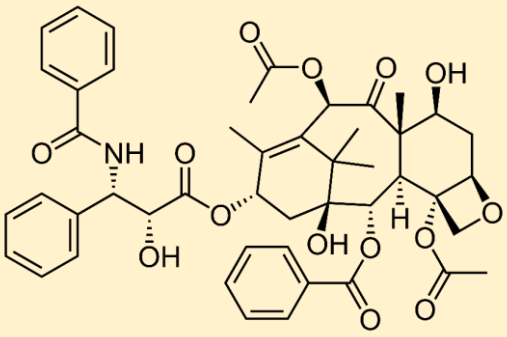
The Pacific Yew Tree

By now, the NCI realized they did not have the time or the resources to supply the nation's ovarian cancer patients with paclitaxel, so they contracted everything to Bristol-Myers Squibb (BMS) in a very controversial deal because the NIH failed to collect the proper monies for Taxol

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

from BMS adequately and gave BMS a virtual monopoly on the drug. BMS invested over $ 1 billion in clinical studies to demonstrate the safety and efficacy of its products. In 1992, Taxol was approved by the FDA. In an interesting role reversal, in 1990, BMS applied to have Taxol as the brand name of the drug and paclitaxel as the generic name. Critics argued that something in the public domain could not be trademarked for 20 years, but after a legal battle, BMS won the right to sell the drug under the name Taxol. It wasn't until 1994 that a total synthetic route for Taxol was elucidated. Taxol is still used to treat ovarian and lung cancer, and in 2000, had peak sales of $1.6 billion [56].

Picture showing how the Pacific Yew tree bark is peeled (top), which is then processed to make Taxol (below)



Taxol is formulated in a mixture of castor oil and alcohol. A version of Taxol dissolved in polyethylene glycol called Taxotere (docetaxol) was approved in 1996.

Image courtesy of Wikipedia.

Following Taxol, other cancer drugs soon followed [57]. Chemists discovered that minor modifications to Taxol's chemical structure could create other anticancer drugs. These drugs were called taxanes, and the most prominent of these was docetaxel, which was approved in 1996. At the height of the NCI's plant collection program, China sent over extracts from *Camptotheca acuminate*, a deciduous tree, which was subsequently found to have anticancer activity. The active ingredient was identified in 1966 as camptothecin. By itself, camptothecin had weak clinical activity; however, changes to its structure ultimately led to the development of topotecan and irinotecan, both of which were approved in 1996. Another plant, *Catharanthus roseus*, more commonly known as the Madagascar periwinkle, has led to the creation of a family of drugs known as the vinca alkaloids. Although plant extracts had long been known to be disinfectants and could be used to treat hypertension and diabetes, the plant sap was poisonous. Soon, the sap components were identified (more than 70), and two components ultimately became vincristine and vinblastine, both approved in 1963.

Taking its cue from the NCI program, Italian researchers began examining soil microbes for their potential antibiotic or anticancer activity. One sample, collected near the Castel del Monte, a 13th-century castle in southeast Italy, contained a new strain of Streptomyces bacterium that produced a bright red pigment with good anticancer activity against various mouse tumors when isolated. Independently, a group of French researchers also discovered the same compound. The two groups named the new compound daunorubicin, derived from "dauni" (where the compound was isolated) and "rub" (for the color red). But in 1967, daunorubicin was shown to cause cardiac toxicity, and its development was halted. The Italian team slightly modified daunorubicin's structure and found the new compound more potent and active against various tumors. This drug was called doxorubicin, and it was approved in 1974. Other drugs, such as cisplatin and chlorambucil, were shown to be effective anticancer agents in the 1970s and, together with other new cancer drugs, marked a turning point in cancer treatment.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Another significant event was the treatment of cancer. Before the 1960s, cancer patients were treated with a single drug, and if that drug didn't work, they were treated with another drug, which continued until the patient went into remission, was cured, or died. In the 1960s, cancer researchers started to show that combinations of drugs could be more effective than their individual components. Researchers found that the drug cocktail called VAMP, which stands for vincristine, methotrexate (formerly known as amethopterin), 6-mercaptopurine, and prednisone, was highly effective in treating childhood leukemia, increasing the cure rate to 60%. Then, researchers started studying combination therapy in solid tumors and obtained similar results. Today, few new cancer drugs are developed under the single-drug paradigm. Today, most new cancer drugs are developed as combination therapy, and it is almost universal to see package inserts that say the drug should be given in combination with another drug.

Castel del Monte in Apulia, Italy, is the soil site where a new strain of Streptomyces was discovered. The bacterium produces a bright red pigment that ultimately led to the discovery of daunorubicin and, shortly after that, doxorubicin.

And while the 1970s saw the introduction of some other major drug classes, like the histamine H2-blockers for the treatment of peptic ulcers, which includes Tagamet (cimetidine), and the non-steroidal anti-inflammatory drugs to treat pain and inflammation, like Motrin (ibuprofen) and Naprosyn (naproxen), the other major event that would have an impact and is still being felt today is the introduction of biologic drugs. In 1976, the world's first biotechnology company, Genentech (which stood for **GEN**etic **EN**gineering **TECH**nology) was founded, on the back of a napkin as legend has it, and they set out to prove that DNA could be recombined from different species and manipulated into making drugs [58].

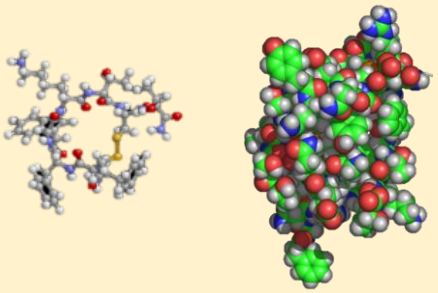
In the late 1970s there was a race to see who could synthesize human insulin first. Many different groups were trying to solve the problem. But synthesizing insulin soon proved to be like the Mount Everest of proteins. While Genentech researchers thought they had a solution, they tackled something less challenging to prove their concept. And for that, they tackled somatostatin. Also known as growth hormone-inhibiting hormone, somatostatin regulates growth and inhibits insulin and glucagon secretion. Whereas insulin has 51 amino acids, somatostatin has only 14. In 1976, taking advantage of advances in molecular biology that had started in the 1950s when Watson, Crick, and Franklin discovered the structure of DNA, Genentech researchers fused the gene for human somatostatin with the gene for *E. coli* beta-galactosidase, i.e., gene-splicing, creating a chimeric plasmid DNA that when inserted back into the bacteria tricks the recombinant bacteria into producing the human protein.

The Founders of Genentech in 1976

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Nothing happened when Genentech researchers first achieved this. They were unable to detect any protein in the cell culture. Upon considering this, they hypothesized that the bacteria were metabolizing the newly synthesized somatostatin as soon as it was formed. To circumvent this, they developed a new gene fusion product, essentially a decoy gene, which linked human somatostatin to another protein that prevented its metabolism. This time, when they experimented, they detected somatostatin in the samples. However, the final product, the first human protein synthesized from a microorganism, was not pure and needed to be cleaved by cyanogen bromide to create the pure somatostatin hormone.

This gave Genentech researchers the confidence they needed to tackle insulin, and soon, a flurry of activity followed. Less than a year later, working with the City of Hope National Medical Center, they repeated their success with insulin, but this time, post-synthesis cleavage was not needed – the protein was fully functional. Clinical trials of the product began in 1980, and in 1982, it was launched as Humulin. At the same time, researchers at Genentech created growth hormone; in 1982, tissue plasminogen activator was produced; and in 1984, Factor VIII for blood clotting was produced. All these products were eventually marketed to patients as drugs.



Top Left: Space-filling model of somatostatin, also called growth hormone inhibiting hormone, a 14-amino acid protein that regulates cell growth and modulates the endocrine system. In 1978, somatostatin was the first protein drug produced using recombinant DNA technology, ushering in the era of biologic drugs. Genentech researchers built the gene to create somatostatin manually, one nucleotide at a time, triplet by triplet.

Top Right: Space-filling model of insulin, a 51-amino acid protein secreted from the pancreas and important for metabolizing carbohydrates, fats, and proteins. In 1982, recombinant DNA-derived insulin was licensed from Genentech. The first marketed biotechnology drug was Humulin, sold by Eli Lilly.

Legend: carbon is green, oxygen red, nitrogen blue, and hydrogen grey. Proteins are not drawn to scale.

The 1980s to 2000s – Rise of Biotechnology

Figure 6 presents a timeline of drugs approved from 1980 to 2000. Genentech's synthesis of insulin and subsequent development of recombinant proteins marks the latest phase in drug discovery – the rise of biotechnology-derived drugs. Genentech was not the only company developing biotechnology-derived drugs. Genentech's proteins were based on chimeric plasmids, where the gene of interest was fused with a bacterium. Other companies were exploring alternative methods and proteins, particularly antibodies. B-cells are the cells in the body that produce antibodies in response to antigens, and there was a theory that these antibodies could be developed into effective drugs against specific targets in the body. But how do you make antibodies of high enough purity, yield, and affinity to manufacture them? Cesar Milstein and Georges Kohler solved this problem in 1975 (they were ultimately awarded the Nobel Prize in 1984).

Their idea was to inject the protein of interest into a mouse so that it would mount an immune response and begin producing antibodies against the protein [59]. The mouse's spleen would then be isolated and fused with myeloma cells, creating a hybridoma, a fused cell of spleen and myeloma origin, which could synthesize antibodies indefinitely under the right conditions. Usually, however, antibodies against the protein are 'polyclonal' in nature, which means that many different B-cells are making antibodies to the protein, but that each B-cell is creating an antibody that binds to a different part of the protein (these are called epitopes). Not all polyclonal antibodies bind to the protein with the same affinity or "tightness;" some have weak binding, some have strong binding. The next step

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

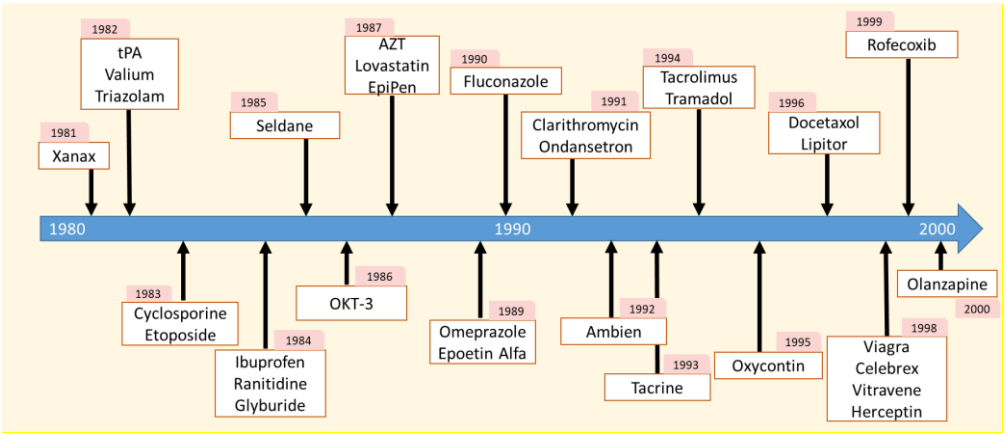


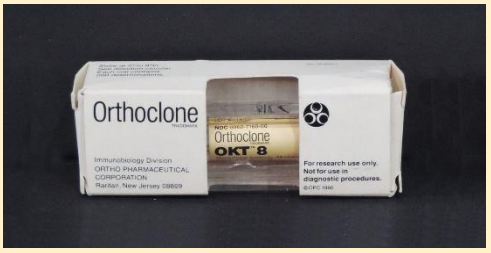
Figure 6. Timeline of important drugs approved from 1980 to 2000.

is to separate the polyclonal antibodies, focusing on those with high affinity towards the epitope of interest. These polyclonal antibodies could then be separated based on their specificity, or how specific the antibody is, towards the protein compared to other proteins. These processes eliminate antibodies with low affinity and specificity to the protein of interest, yielding a "monoclonal antibody" with high affinity and specificity towards a single epitope on the protein of interest. Today, alternatives to hybridomas exist, such as phage display technology, which utilizes molecular biology techniques like the polymerase chain reaction (PCR) to increase the amount of heavy and light chains of the antibody.

In 1986, Ortho Biotech released the first monoclonal antibody created using hybridoma technology - Orthoclone OKT3 (muromomab-CD3) to treat acute kidney transplant rejection. OKT3 is an IgG2a antibody that binds to and blocks the CD3 receptor, which normally activates cytotoxic T cells. Within a few minutes of administration, CD3+ T-cells virtually disappear from plasma, making it a useful immunosuppressive drug. One problem with OKT3 is that the IgG framework upon which it was built was derived from mouse spleen cells. With repeated administration, the human body recognizes OKT3 as a foreign protein. It begins mounting an immune response against the drug, leading to the formation of HAMA (human anti-mouse antibody). When HAMA antibodies are formed, the efficacy of OKT3 is reduced. To prevent the formation of an immune response against antibodies, future antibodies would be built on increasingly human-like frameworks, with currently developed monoclonal antibodies being fully human in structure.

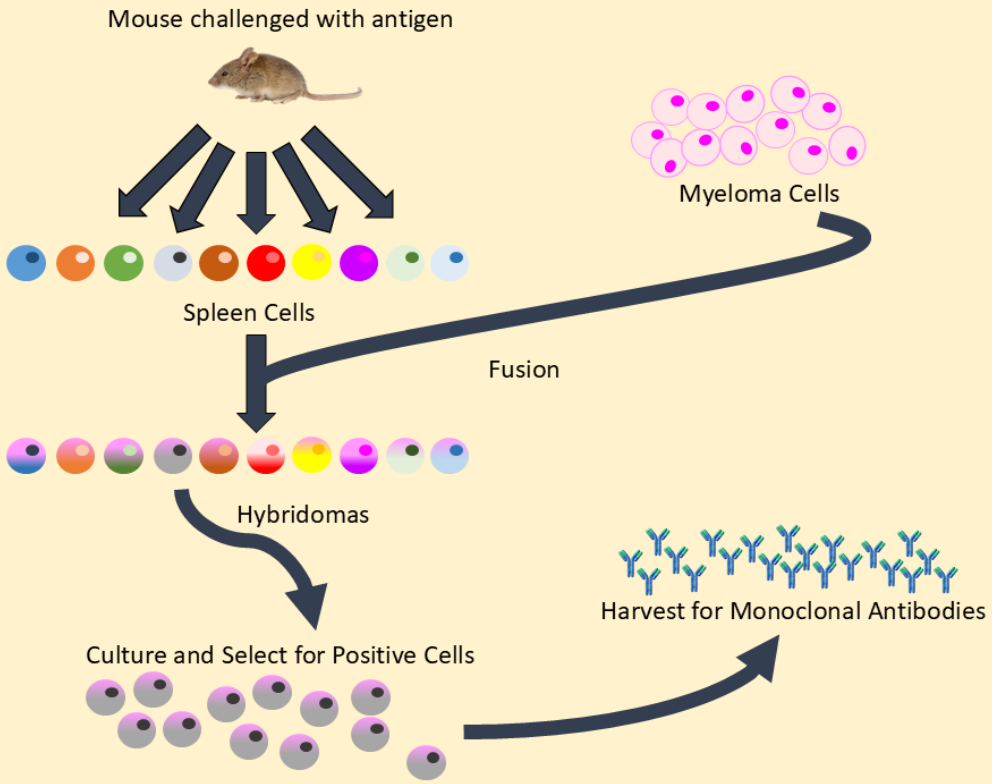
Since OKT3's release, monoclonal antibodies have exploded onto the scene. What was once a specialty area inhabited by biotechnology companies is now Big Pharma's mainstay. As of 2022, 162 monoclonal antibodies have been approved for therapeutic use worldwide, and they are active against more than 90 different targets [60]. At the same time, the global market for monoclonal antibodies was estimated to reach more than $600 billion by 2032 [61]. Humira alone, the largest-selling monoclonal antibody on the market at the time, had sales of $21 billion in 2022, followed by Keytruda and Stelara with sales of $21 billion and $10 billion, respectively.

In 1986, Orthoclone OKT-3 (muromab-CD3) was the first approved monoclonal antibody and was developed using murine hybridoma technology.



Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Hybridoma Technique for Antibody Production



Animals, particularly mice, when injected with a protein of interest, will generate an antibody response. The B cells that produce antibodies to the injected protein are then isolated from the animal's spleen and fused with myeloma cells that have been modified to express a selectable marker, creating a hybridoma. A hybridoma is needed because spleen cells do not grow in culture. Myeloma cells do not produce antibodies but do replicate in culture. Hence, a hybridoma utilizes both cell lines to create an essentially immortal, replicating cell line that produces antibodies. A trick is used to separate the hybridoma cells from the originator and myeloma cells.

Usually, for cells to replicate, they must first replicate their DNA. Cells can make DNA by two pathways: the de novo pathway and the salvage pathway. When the de novo pathway is blocked, cells use the salvage pathway, which requires the nucleotides hypoxanthine and thymidine. The trick is to use myeloma cells that lack an enzyme called HGPRT. When this enzyme is absent, the salvage pathway of DNA synthesis will not occur. Also, the media the cells are in contains aminopterin, which blocks the de novo synthesis pathway. Hence, DNA synthesis will not occur in the myeloma cells because both pathways are blocked. Hybridomas, on the other hand, can replicate because they inherit a functional HGPRT from spleen cells. With hypoxanthine and thymidine also present, they can replicate their DNA through the salvage pathway. Hence, the spleen and myeloma cells will die, leaving functional hybridoma cells that can produce the desired antibodies.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drug discovery from the 1980s onward harnessed the power of chemistry, but the biotechnology revolution harnessed the power of biology. To say that the use of recombinant DNA methods to produce biotechnology drugs was a medical revolution is not such a stretch. Small molecule drugs, while often potent, can target many different receptors in the body – they are not very specific and, consequently, can have many side effects. In contrast, biotechnology-derived drugs are highly specific and have few off-target side effects. The side effects they do have are often related to their route of administration and exaggerated pharmacological effects at their intended target. Due to the challenges of producing biological drugs, this has proven a disincentive to generic competition for most generic manufacturers. Thus, biological drugs are as common in pharmaceutical companies today as their small-molecule counterparts.

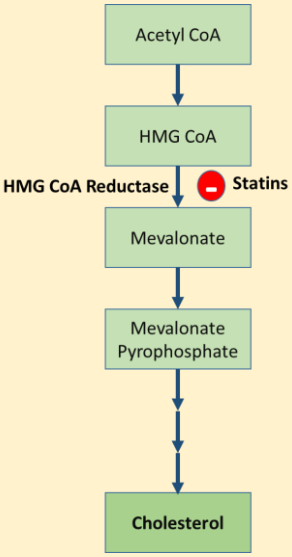
The political climate in the 1980s was toxic as the government, patients, and researchers were often at odds in the search for medications to treat HIV. AIDS activist groups like ACT UP charged the FDA and the National Institutes of Health with taking too long to develop approved medications for the treatment of HIV. As a result of their activism and public outcry, the FDA reevaluated its approval process for drugs for life-threatening illnesses and established a new basis for approval, utilizing a surrogate marker that predicts the actual clinical endpoint. In this case, the FDA allowed CD4 counts and viral load as surrogate markers for survival in HIV patients. Because death may not occur for many years, it may take a long time before the effect of a drug on survival can be objectively assessed. A surrogate marker is a short-term marker that indicates the long-term effect. Changes in a surrogate marker can quickly show whether a drug may affect the clinical endpoint.

The 1980s would not be complete without at least mentioning that it was during this time that Human Immunodeficiency Virus (HIV) drugs became available, as did the first statin drug for the treatment of hypercholesterolemia. The story of AIDS, HIV, and the drugs used to treat HIV could be an entire book in itself. Although today, HIV is a largely manageable disease, that was certainly not the case 30 years ago. From 1981 to 1987, the number of new HIV/AIDS cases doubled every year, such that by 1987 there were roughly 40,000 deaths and 50,000 new cases. The politics at the time were unimaginable, with protesters, the government, and researchers often at odds. But in 1987, Retrovir (zidovudine, also known as azidothymidine or more commonly as AZT) was the first drug approved to treat HIV [62]. It was not a great drug; the side effects were hard to tolerate, and it prolonged life by about a year, but it was better than nothing, and it offered the promise, if not with this drug, then with future drugs, that the virus could be contained. It would not be for another ten years before new drugs and drug cocktails (HAART therapy) turned HIV from a death sentence to a long-term illness.

The other major class of drugs introduced in the 1980s was statin drugs to treat hypercholesterolemia; these were not biologics but small molecules. Cholesterol is a fatty compound necessary for human life. It is a major component of cell membranes, forms the foundation for many steroid hormones, like estrogen and testosterone, and may have anti-inflammatory effects. However, excessive cholesterol can form plaques on the inner walls of arteries, leading to arterial hardening and making it more difficult for blood to circulate throughout the body. These plaques can break off and form blood clots that could cause a stroke or heart attack. By the mid-1980s, it was generally accepted that lowering cholesterol would reduce the frequency of cardiac events. HMG-CoA reductase was identified as a valuable target for lowering cholesterol in the 1970s, when natural products that inhibited its activity in the liver were found to reduce plasma cholesterol levels in animals.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

The 1980s also saw the introduction of cytokine therapy. Cytokines are substances secreted by T-cells that have an immunomodulatory effect on other immune cells [63]. Cytokine therapy aims to administer exogenous cytokines to directly target cancer cells or modulate the function of other cells. The first approved cytokine was interferon-α (Intron), approved in 1986 for hairy cell leukemia. Others quickly followed. Granulocyte colony-stimulating factor (G-CSF, Neupogen) was approved in 1991 to stimulate the growth of white blood cells. Interleukin-2 (Proleukin) was approved in 1992 for the treatment of renal cell carcinoma. Many other cytokines are in development, including vascular endothelial growth factor (VEGF) and various other interleukins (IL-12, IL-6, IL-13, etc.). Cytokine therapy, while effective at treating cancer, has adverse events that are frequently intolerable and have short half-lives, requiring large doses to achieve effect. Many current cytokine programs are designed to overcome these challenges, e.g., pegylated formulations.



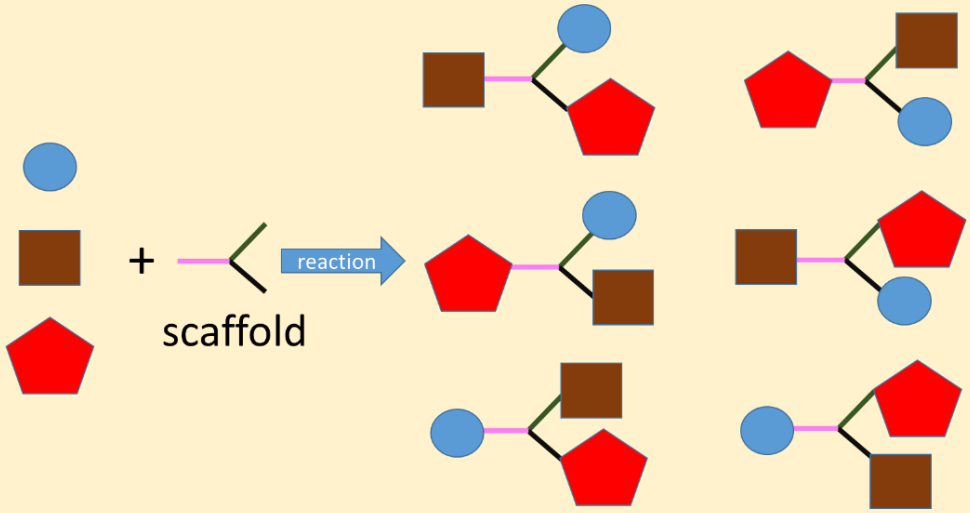
The Cholesterol Biosynthesis Pathway. Statins block the conversion from 3-hydroxy-3-methyl-glutaryl-coenzyme, or HMG CoA for short, to mevalonate by blocking the enzyme HMG CoA reductase. This step is rate-limiting in cholesterol synthesis. Blocking this step ultimately halts cholesterol synthesis and lowers blood lipid levels.

In 1978, Merck researchers discovered a potent inhibitor of HMG Co-A reductase in the fermentation broth of the fungus *Aspergillus terreus*. They named their discovery mevinolin, which was later given the INN name lovastatin. With the INN name lovastatin, the statin name was born. In clinical trials, lovastatin led to a significant decrease in LDL cholesterol, a slight decrease in plasma triglycerides, and a modest increase in HDL cholesterol. In 1987, lovastatin was approved and marketed as Mevacor. Shortly after that, Zocor (simvastatin) and Pravachol (pravastatin) were approved in 1991, Lescol (fluvastatin) in 1993, and Lipitor (atorvastatin) in 1996. In 2009, when Lipitor sales reached their peak, the total statin market was approximately $20 billion. Today, because of the launch of generic drugs, the statin market is not nearly as high. Still, it has been estimated that statin use in the United States was 92 million users in 2018-2019, representing more than a quarter of the population [64]!

The 1990s also marked a paradigm shift in the way drugs were discovered and the end to some of the great stories of how drugs were accidentally discovered. Before then, drug discovery research was chemocentric. Starting from a particular compound or compound class, a series of chemical modifications was made and tested for pharmacologic activity using in vitro or in vivo animal models. Such tests for activity are called phenotypic screens. The pharmacologist was often unsure what the molecule would do, so the screens were designed to test for various effects. Basically, the philosophy was, 'I have a drug; let me find an indication for it.'

In the 1990s, several significant developments took place. First, with the advent of molecular biology, it became possible to clone the receptors to which drugs bind from their corresponding genes. Pure receptors enabled the synthesis of drugs that could specifically target them. Chemists could isolate the active site, determine its amino acid composition, perform X-ray crystallography to determine its 3D structure, and then synthesize a drug that fits into the pocket. The magic bullet hypothesis could be realized. Second, chemistry developed what is called combinatorial chemistry. Starting from a chemical scaffold, a series of analogs would be rapidly developed automatically, without the level of manual effort previously required. Hence, researchers could quickly develop hundreds to thousands of analogs from a single chemical backbone.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



The Combinatorial Chemistry Paradigm. Combinatorial chemistry is a powerful technique that chemists use to rapidly generate thousands of different compounds. The scaffold is the part of the molecule that remains constant. The shapes represent the possible functional groups that can be added to the scaffold, in all possible combinations. As one might expect, the possible "chemical space" increases with scaffold size and the number of functional groups.

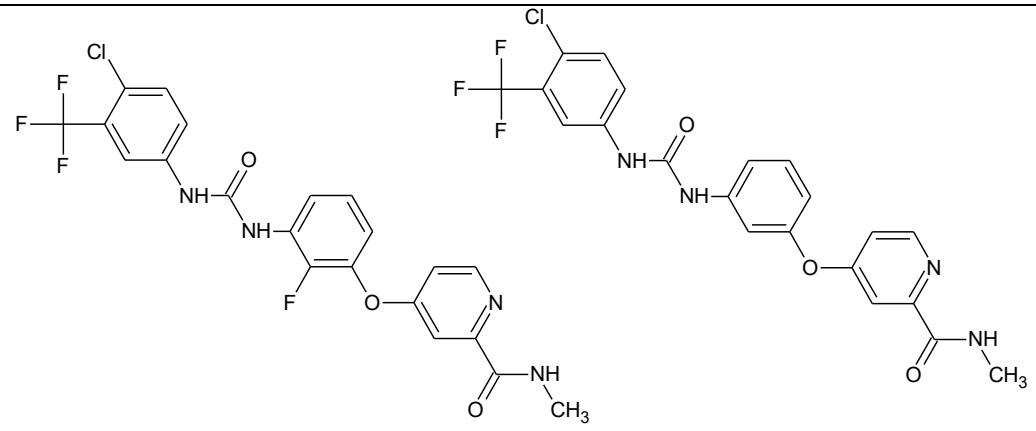


Figure 7. Example of a me-too drug. sorafenib (on the right) was approved in 2007 for the treatment of hepatocellular carcinoma. Regorafenib (on the left) was approved for the same indication in 2012; however, the approval was updated in 2017. Sorafenib improved survival by about 2 months compared to placebo. Regorafenib improved survival by the same amount. The primary structural difference between the two is the presence of a fluorine atom on regorafenib. Regorafenib is considered a me-too drug [65]. While competitors develop most me-too drugs, this particular me-too was a life cycle strategy because Bayer developed both drugs, which later went on and had sales exceeding $1 billion in 2023.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Gone would be the days when the chemist would toil in the lab, slowly synthesizing hundreds to thousands of compounds during their careers. Instead, robots could synthesize the same numbers in a fraction of the time. And lastly, high-throughput screening (HTS) has matured. Using pure cloned receptors, scientists could rapidly determine which analogs from the combinatorial chemistry screen had the highest binding affinity to the receptor.

Alternatively, scientists could screen what are called compound libraries, vast sets of analogs with different scaffolds, to find those compounds with the best binding properties to a particular receptor. Using these three advances, drug discovery transitioned from a chemocentric to a target-based approach – once the target is identified, the most effective drug that binds to it can be found. Those compounds with a high degree of binding were called 'hits.' Starting from these hits, the chemist would characterize the binding determinants, whether a free hydroxyl group or a long lipid chain, and then attempt to optimize the binding even further. This process is called 'hit to lead optimization' [66]. From a library of 5,000 to 10,000 compounds, only a handful would be advanced to the lead optimization stage.

Although the 1990s had some notable discoveries, by and large, it can be thought of as the me-too decade. Me-too drugs are pharmacologically active compounds that are structurally related to already approved drugs in the same therapeutic class and differ in some respect from those drugs, e.g., by having different receptor binding affinity, different pharmacokinetics, different adverse event profiles, etc. [67]. In one study, which has since been pulled off the internet, nearly two-thirds of all approved drugs from 1989 to 2000 were copycat, or me-too drugs, of ones already on the market, rather than new drugs with innovative mechanisms of action [68]. in Figure 7 shows an example of a me-too drug, which compares the structure of sorafenib (the originator) against a later competitor, regorafenib. The only structural difference between the two is the presence of a fluorine atom in regorafenib. Both drugs had nearly identical improvements in survival compared to placebo.

One notable development in the 1990s and 2000s was the emergence of biologic drugs. Biologics, particularly monoclonal antibodies, started taking center stage, demonstrating how powerful these drugs could be. The methods used to create monoclonal antibodies have resulted in more human-like products. In the 1990s, 30% of monoclonal antibodies were based on mouse binding domains; however, by 2008, this number had decreased to 7%, while the percentage of fully human immunoglobulins had increased to 45%. The 1990s also saw the introduction of recombinant enzymes as drugs for humans. Ceredase (alglucerase) was a modified form of β-glucocerebrosidase, a lysosomal enzyme responsible for glycolipid metabolism in the cell, which is non-functioning in patients with Gaucher's disease. Patients with Gaucher's disease have a build-up of glucocerebroside in the spleen, liver, and macrophages that can lead to enlarged liver and spleen, neurological defects, severe joint pain, and osteoporosis. Ceredase was initially manufactured from human placental tissue and, in 1991, was the first approved enzyme replacement therapy. Later, a recombinant DNA version

The 1990s would not be complete without discussing Viagra, probably one of the last drugs discovered by accident. Sildenafil was originally tested as a cardiovascular drug for angina, but in the early clinical trials in humans (which at that time were

mostly men), an unusual side effect occurred. When the nurses went in to check on them, many of them were on their stomachs and did not want to roll over because they all had erections and were embarrassed. It appears that not only were blood vessels in the heart dilated, but blood vessels in the penis were as well. The rest is history. Viagra went on to create an entirely new erectile dysfunction drug category and a host of me-too products, with peak Viagra sales exceeding $2 billion in 2012. Pfizer later went back and tested sildenafil for its original indication, and in 2005, the drug was approved as a treatment for pulmonary arterial hypertension. To avoid confusion, Pfizer markets the product as Revatio.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

that did not need human placental to be made was approved in 1994 and sold as Cerezyme.

Biotechnology-derived products, such as recombinant enzymes, are first produced in cells grown in the laboratory and then transferred to bioreactors (shown above), which provide all the necessary nutrients for cell growth. Some of these may be as large as 20,000 liters. The desired protein is then isolated from the cells and medium in the bioreactor and filled into vials for human administration.

The 1990s also saw the continuing rise of cancer treatments, particularly a new class of agents – targeted chemotherapy. Cancer drugs developed before the 1990s were considered cytotoxic agents in that they indiscriminately killed cells, both cancer and noncancerous cells, but particularly liked to kill the rapidly dividing cells. This led to adverse events like nausea, ulcerations of the gastrointestinal tract, diarrhea, bone marrow suppression, and hair loss. In the 1980s, British researchers discovered that the ABL cancer gene was located on chromosome 9. While this may seem mundane, it wasn't because 95% of all patients with chronic myelogenous leukemia (CML), a type of cancer of the white blood cells, have a chromosomal translocation called the Philadelphia chromosome, where the ABL1 part of chromosome 9 breaks off and sticks to the BCR part of chromosome 22, creating a fusion gene called BCR-ABL1. This fusion gene encodes a BCR-ABL mutant tyrosine kinase protein that remains active, leading to uncontrolled cell division.

Now that researchers knew the target, they could find a drug to inhibit it. Scientists at the Dana-Farber Cancer Institute in the United States collaborated and screened hundreds of compounds to find one that would inhibit this protein. They found one, which was called STI-571, and it worked in test tubes and animals [69]. When given to CML patients and to patients with another type of leukemia that sometimes has Philadelphia chromosome aberrations, acute lymphoblastic leukemia (ALL), the results were amazing - 55% of CML patients showed a positive effect, with four patients showing a complete cure. Of those who did not show a complete cure, their illness went into remission for ~3 months to 1 year. Of the 20 ALL patients, 14 (70%) responded, with 4 achieving complete remission. Imatinib also had an impressive safety profile and was much less toxic than cytotoxic therapies.

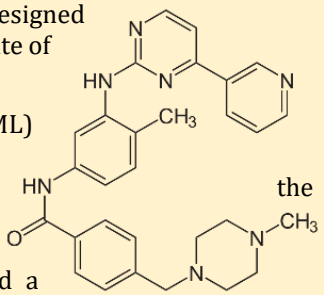
By one estimate, almost 1 in 4 of all drugs approved for use between 1938 and 2012 were discovered by chance!

STI-571 was renamed imatinib and, in 2001, was approved for use and marketed as Gleevec. Gleevec was huge – it is considered the first targeted cancer therapeutic, specifically targeting a particular cancerous protein. Later, Japanese researchers showed that imatinib was also useful in treating gastrointestinal stromal tumors, which develop in the connective tissue of the gastrointestinal tract. One problem with imatinib is that patients may become tolerant to its effects and develop resistance. In 2006, Sprycel (dasatinib) was approved for the treatment of chronic myeloid leukemia (CML) in patients who are resistant to imatinib.

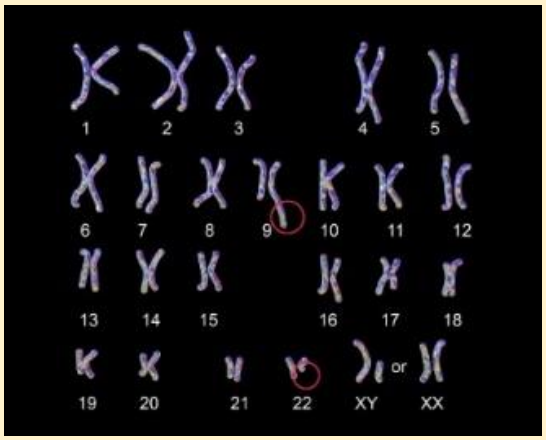
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Discovery of Gleevec, a Rationally Designed Cancer Drug

Gleevec (imatinib mesylate) is the poster child of rationally designed drugs and may be single-handedly responsible for the current state of targeted drug development in the pharmaceutical industry.



About 10,000 new cases of chronic myelogenous leukemia (CML) occur every year in the US and about 1,300 people die as a result [70]. Roughly 1 in 526 people will get CML in their lifetime. In the 1950s, when scientists were trying to understand how genetic alterations caused cancer, Peter Nowell and his graduate student, David Hungerford, discovered that CML patients had a small, atypical chromosome in their cancer cells, which was later coined the "Philadelphia chromosome," after the city in which it was discovered. Surprisingly, no other cancer has such a consistent chromosomal change. Twenty years later, Janet Rowley discovered that the Philadelphia chromosome was the result of the fusion of two distinct pieces of DNA. In this translocated chromosome, the long arm of Chromosome 22 is substituted with the long arm of Chromosome 9.



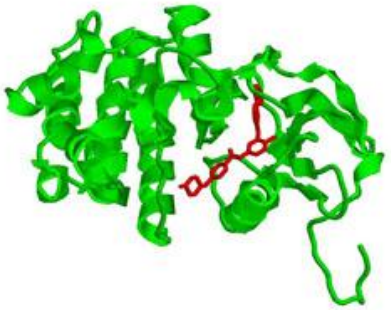
Ten years later, researchers identified the translocated genes on the Philadelphia chromosome. This fusion product was called bcr-abl, and later work showed that bcr-abl produces an overproduction of a tyrosine kinase protein. Kinases typically phosphorylate proteins, thereby activating them. Tyrosine kinases are involved in cell signaling, causing cells to divide. In CML patients, 95% of whom have bcr-abl, this results in an overproduction of immature white blood cells, with counts 10- to 25-times higher than normal. What makes these findings interesting is that CML is caused by the overproduction of a single protein, thereby setting the stage for the development of imatinib.

Ciba-Geigy (later acquired by Novartis) was developing inhibitors that would block the ATP-binding pocket of tyrosine kinases, thereby inhibiting their activity. In cell culture, one of these inhibitors, CGP 57148, later renamed STI-571 and eventually imatinib, demonstrated more than 90% inhibition of tumor growth without affecting normal cells. In Phase 1 trials with CML patients, complete regression of tumor size was observed in 53 of 54 patients (an unprecedented 98% response rate) within 4 weeks of treatment [71]. In 2001, the drug was approved by the FDA and subsequently demonstrated effectiveness in several types of cancer. In a sense, imatinib started the targeted therapy revolution in cancer by validating the "magic bullet" hypothesis of cancer therapy. It may have led to false expectations because the scientists Ciba-Geigy/Novartis made it look easy.

Figure courtesy of Wikipedia Creative Commons.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Imatinib opened the door for a whole host of targeted therapeutics, drugs that specifically targeted the cancers they were treating while sparing non-cancerous cells. Once imatinib demonstrated that it could inhibit the products of these oncogenes, such as BCR-ABL, researchers began examining other cancers and their gene products, leading them to focus on tyrosine kinases, like BCR-ABL, which phosphorylate proteins at tyrosine residues. Protein phosphorylation generally activates that protein. Tyrosine kinases mediate a host of roles in the body, including cell growth, differentiation, metabolism, and apoptosis [72]. Studies have shown that mutation, overexpression, and overstimulation of these kinases may lead to cancer.



Cartoon of the binding of Imatinib (red) to the Mutant ABL Fusion Protein Tyrosine Kinase (green)

The floodgates opened as researchers sought compounds that would inhibit various tyrosine kinases. They developed drugs such as Tarceva (erlotinib), Sutent (sunitinib), and Votrient (pazopanib), among others, which led to the creation of an entire class of drugs known as tyrosine kinase inhibitors (TKIs). However, researchers did not stop at small-molecule inhibitors. They also began to focus on monoclonal antibodies targeting various tyrosine kinases for cancer treatment, which had the potential to be highly specific in blocking their targets. An example of this is Herceptin (trastuzumab), which was approved in 1998 for the treatment of breast cancer [73]. Trastuzumab is a monoclonal antibody that blocks the HER2 receptor, which is overexpressed in breast cancer, and inhibits the activity of several growth factors. Without HER2 overexpression, trastuzumab has no effect, and patients must be tested for HER2 overexpression before they can be treated with the drug. Today, research has progressed beyond cytotoxic drugs that are not specific to cancer cells and now focuses exclusively on targeted therapies that target particular proteins or other specific targets. Furthermore, TKIs have expanded beyond cancer to other indications, such as autoimmune and inflammatory diseases, with the approval of several JAK (another tyrosine kinase) inhibitors, including Xeljanz (tofacitinib). As of 2023, over 50 targeted kinase inhibitors (TKIs) have been approved by the FDA, and the global market for TKIs is expected to exceed $85 billion by 2028.

Closing out the 1990s marked the first clinical trials for gene therapy. Ever since the 1960s, researchers have believed that DNA could be used to treat various diseases; however, direct DNA injection is short-lived because blood enzymes rapidly degrade it. Two types of gene therapies would soon arise: in vivo and ex vivo. In ex vivo gene therapy, cells are removed from the patient, modified, and reintroduced using a vector. In vivo gene therapy involves the exogenous administration of a gene, packaged in a vector, into the body via routes such as intramuscular or intravenous injection. In 1990, the first clinical trial of an ex vivo gene therapy was performed. Ashanti DeSilva was 2 days old when she developed an infection, the first of many, before she was eventually diagnosed with a genetic condition where her body made defective forms of the adenosine deaminase (ADA), an enzyme that breaks down adenosine. Blood was removed from the body, the cells were infused with a functional ADA gene, and then injected back into her bloodstream. The trial was a success, Ashanti was cured, and today she is a genetic counselor. Other trials were started, and patients were dosed until 1999 when Jesse Gelsinger was treated for another genetic enzyme deficiency, ornithine transcarbamylase. By then, he was the 18th person treated with ex vivo gene therapy. This time, however, Jesse went into multiorgan failure and died. He was 18 years old. The FDA immediately suspended all gene therapy trials. Lawsuits were filed. Much was made of the informed consent process and its lack of clarity. It took years before the next gene therapy trial could be started. Jesse's trial was a major factor in the development of in vivo gene therapy and the use of modified adenovirus vectors as gene-delivery vehicles.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

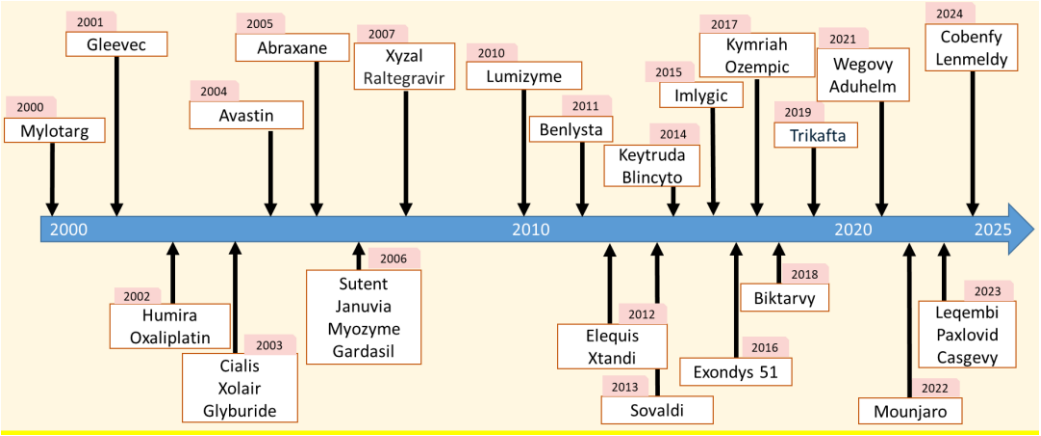


Figure 8. Timeline of important drugs approved from 2000 to the 2020s.

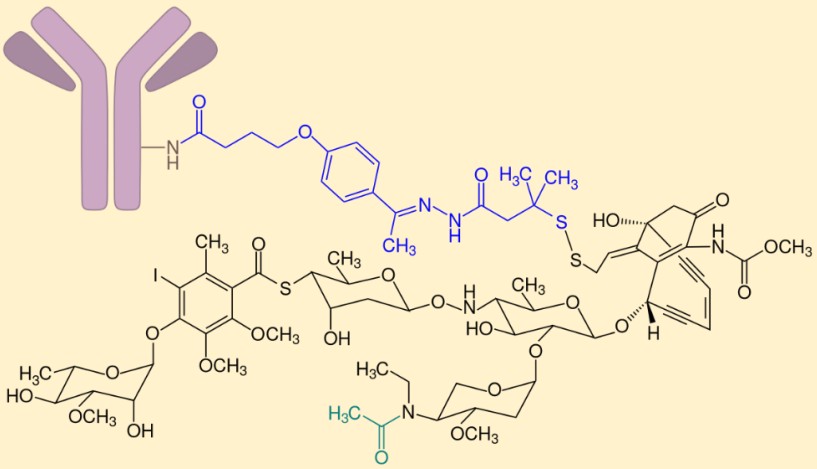
The 2000s to 2020s

Figure 8 presents a timeline of important drug approvals from 2000 to the 2020s (up to 2025). One antibody that deserves particular attention was the approval of Humira (adalimumab) for the treatment of rheumatoid arthritis (RA) in 2002. Tumor Necrosis Factor (TNF-α) is known to be an important mediator of inflammation. Too much TNF in the blood can lead to persistent tissue damage and pain. Furthermore, as mentioned earlier, hybridomas from mice were poor drugs due to the formation of antibodies against the monoclonal antibody. In the late 1980s, Greg Winter developed an alternative to hybridomas, known as phage technology, at the Medical Research Council's Laboratory of Molecular Biology in the UK. Bacteriophages, or phages for short, are viruses that infect and replicate only in bacteria, not humans. The genomes of these phages are modified by inserting a fragment of a specific antibody gene. The resulting antibody fragment is then expressed on the surface of the phage capsid, which can later be used to identify particular phages for purification. The phage then infects Escherichia Coli bacteria, which can generate the antibody of interest in large amounts. Winter was awarded the Nobel Prize in Chemistry in 2018 for this discovery.

Cambridge Antibody Technology (CAT) in the UK was established to leverage this method and develop a pipeline of monoclonal antibodies that could be further refined into drugs. In 1993, BASF Pharma, in collaboration with CAT, developed a monoclonal antibody to TNF-α called DE27. In 1998, human trials demonstrated that DE27 was effective in treating rheumatoid arthritis. In 2001, Abbott acquired BASF and DE27. In 2002, adalimumab, given the generic name DE27, was approved by the FDA under the brand-name Humira, marking the first drug approved using phage display technology. Humira would become the best-selling drug of all time, earning more than $200 billion globally, with over 400,000 patients being treated with the drug [74].

For the most part, the 20th century has been characterized by significant advancements in cancer drug development. The end of the 20th century also saw the combination of a monoclonal antibody with a cytotoxic drug, known as the payload, to treat cancer. These drugs are called antibody-drug conjugates (ADCs). Cytotoxic cancer drugs, as you recall, were the original cancer drugs that are nontargeted and are toxic to all cells in the body to some degree. The idea was that the antibody would target a specific tumor, be internalized by the tumor, release the cytotoxic drug payload within the tumor, and kill it, leaving normal cells untouched. ADCs require a stable linker between the antibody and the payload that releases the payload inside the tumor. In 2000, the first ADC, Mylotarg (gemtuzumab ozogamicin), was approved for the treatment of acute myelogenous leukemia in patients older than 60 years. Mylotarg combined an IgG5 immunoglobulin against CD33, a cell surface protein expressed on certain white blood cells, and calicheamicin, a cytotoxic drug that stops DNA

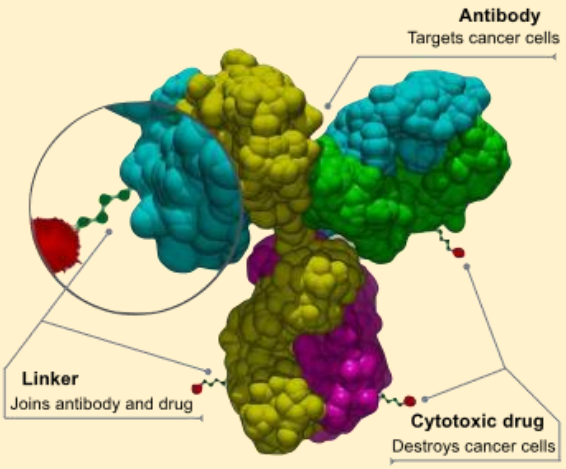
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Mylotarg was the first approved antibody-drug conjugate (ADC) that combines a monoclonal antibody targeting a cell surface receptor on white blood cells. The cytotoxic payload would be released inside the cell, stopping DNA replication and killing the tumor. Mylotarg was removed from the market in 2010 because of safety concerns, but was later reapproved in 2017. Figure reprinted from Wikipedia.

replication. It was thought that targeting the tumor more specifically with the antibody would reduce the payload's toxicity. In the case of Mylotarg, this was not necessarily the case, as a black box warning was added to the label a year later, indicating an increased risk of veno-occlusive disease. In this potentially life-threatening side effect, the small veins of the liver are obstructed, often leading to acute kidney failure. By 2010, after randomized clinical trials showed a high rate of adverse events, Mylotarg was removed from the market. However, Mylotarg was reapproved in 2017 (with a black box warning still there).

In 2011, a year after Mylotarg was removed from the market, Adcetris (brentuximab vedotin) was the second ADC approved, followed by Kadcyla (trastuzumab emtansine) in 2013, both of which have black box warnings. By 2021, 14 ADCs had been approved, all of which carried black-box warnings for the treatment of solid tumors and hematologic malignancies. While the concept of ADCs supports their efficacy, a risk associated with their use does appear to exist, and further research is needed to enhance their safety profile. As of 2022, it is estimated that more than



Antibody-drug conjugates (ADCs) combine a monoclonal antibody with a cytotoxic small molecule, known as a payload. ADCs are taken up into cancer cells, the payload is released, and the tumor cells are killed. Modern ADCs have monoclonal antibodies linked to receptors or smaller antibody fragments. Figure courtesy of Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

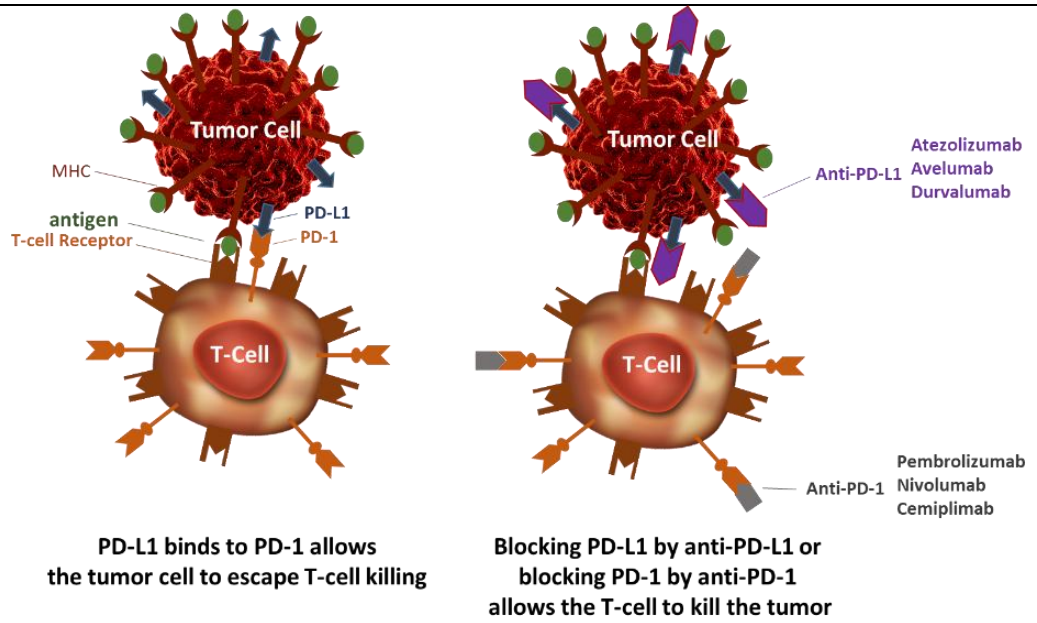


Figure 9. Immune checkpoints are protein-protein interactions between cells that have a stimulatory or inhibitory effect. Tumors evade death when checkpoint proteins, such as PD-1 on T cells, bind to PD-L1 on tumor cells, preventing the T cell from killing the tumor (left). T-cells have several checkpoint proteins, including CTLA-4, PD-1, CD27, GITR, and TIGIT. Checkpoint inhibitors prevent the binding by blocking PD-1 or PD-L1 in this example, allowing the T-cell to kill the tumor (right). Korman, Peggs, and Allison were the first to coin the term "checkpoint blockade" to describe drugs that block receptors on T-cells [76].

100 ADCs are in development, with 14 approved for use [75].

The other significant development in oncology in the early 2000s was the emergence of immuno-oncology (IO), which utilizes the immune system to combat cancer. To say that IO revolutionized oncology is not an understatement. Scientists have been trying to harness the immune system to treat cancer since the 1860s, when Rudolf Virchow, the father of cellular pathology, first identified white blood cells in cancer tissues. Later, it was shown that many cancers are associated with long-term inflammation and that anti-inflammatory drugs reduce the risk of cancer. Finally, in the 1960s, immune cells were demonstrated to protect mice from cancer. Since then, the holy grail of cancer treatment has been to harness the immune system to target and kill cancer cells. The human body has various types of cells, including natural killer cells, killer T cells, helper T cells, and B cells, that effectively kill cancer cells. However, being tricky little devils, cancer cells have developed mechanisms to avoid these immune responses. Cancer biologists have been studying for decades how to overcome these resistance mechanisms.

Today, there are many types of IO drugs, but one class in particular has been checkpoint inhibitors. Immune checkpoint blockade represents the last arm of immunotherapy [77]. One way the body promotes self-tolerance, i.e., the ability to avoid eliciting an immune response against normal cells, is through checkpoint inhibition, which regulates the immune system. These checkpoints are a normal part of the immune system; tumor cells have evolved cell surface proteins that activate immune checkpoints, allowing them to evade destruction. When checkpoint proteins bind, the T cell is prevented from killing the tumor. A class of new drugs called checkpoint inhibitors prevents the binding between checkpoint proteins, allowing T-cells to kill the tumor (Figure 9).

Since the 1980s, it has been known that activation of T-cell receptors alone is insufficient to activate T-cells, and a costimulatory signal is also required. This was identified as the binding of the protein

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

B7 to the cell-surface protein CD28, also on the T-cell. In 1987, a second co-stimulatory protein was discovered in activated cytotoxic T cells, called CTLA-4, which was later identified as a negative regulator of T-cell function. In 1996, James Allison and colleagues demonstrated that mice implanted with tumors and treated with an antibody to CTLA-4 exhibited anti-tumor activity [78]. This spurred researchers to develop a monoclonal antibody for human use. In 2011, Yervoy (ipilimumab) was approved for the treatment of melanoma, which has since been expanded to many different types of cancers. The field then exploded, and soon after, other checkpoint proteins and blockers were approved, including atezolizumab, avelumab, and durvalumab for PD-L1 blockade, and pembrolizumab, nivolumab, and cemiplimab for PD-1 blockade. As of 2022, 9 checkpoint inhibitors have been approved by the FDA, and dozens more are in clinical trials. The pioneering work of Allison and Tasuku Honjo earned them the Nobel Prize in Medicine in 2018.

A new class of drugs that has shown promise for treating various cancers is bispecific antibodies (BsAb). In a normal monoclonal antibody, the two arms, the Fab regions, bind the same antigen. With a bispecific antibody, the Fab regions bind different antigens. This differential binding allows them to act as a kind of bridge between different cells, like a tumor and a T-cell, thereby bringing them closer together so that the T-cell can act against the tumor. The first approved BsAb was Blincyto (blinatumomab) in 2014 for the treatment of relapsed or refractory acute lymphoblastic leukemia, which targets CD19 on B cells and CD3 on T cells. As of 2025, nearly a dozen BsAbs have been approved for various indications, and dozens more are currently in development. Further, many drugs are in development that are modifications to the classic BsAb design; drugs with funny names like BiTE, DART, and TRiKE are being developed to try to improve on BsAb performance.

Other notable advances in oncology in the 2010s were oncolytic viruses and cancer vaccines. Oncolytic viruses utilize genetically modified viruses, such as smallpox or adenovirus (the common cause of respiratory illnesses like the common cold), to infect cancer cells and create a pro-inflammatory tumor microenvironment that boosts natural antitumor immunity [79]. In 2015, talimogene laherparepvec (Imlygic), an engineered herpes simplex virus, was the first oncolytic virus approved and is used to treat melanoma. Other oncolytic viruses approved in the US are nadofaragene firadenovec (Adstiladrin), an engineered adenovirus for treating a particular type of bladder cancer. Cancer vaccines utilize antigens specific to tumors to trigger antitumor T-cell responses [80]. Gardasil was the first approved cancer vaccine, approved in 2006, and is 100% effective in preventing a variety of cancers caused by human papillomavirus, the foremost being cervical cancer. Heplisav-B was approved in 2017 for the prevention of infection caused by hepatitis B, which is a causative factor in liver cancer.

The portfolios of many large companies have changed over the last 20 years, shifting from predominantly small molecules to biologics. For example, small molecules and biologics accounted for 62% and 38% of Pfizer's portfolio in 2011, but those numbers flipped completely by 2019 – 62% for biologics and 38% for small molecules [81].

If it seems like most advances in drug discovery over the last 20 years are in oncology, you may not be wrong to think that. Today, oncology is the major therapeutic area pursued by drug companies [82]. Nearly all major pharmaceutical companies have an oncology division. This is compared to cardiovascular, the dominant therapeutic area in 1995. The top five therapeutic areas in 1995 (cardiovascular, anti-infective, antiulcer, neuroscience, and cholesterol), which accounted for 53% of all therapeutic areas, dropped to just 6% in 2020! Conversely, the top 5 therapeutic areas in 2020 (oncology, immunology, anti-diabetes, anticoagulants, and HIV antivirals) accounted for 62% of total spending, compared to just 11% in 1995. The reason oncology is the top area is the huge profits that can be achieved, given lower approval barriers and lower drug development costs compared to cardiovascular or anti-diabetes drug development.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Aduhelm Controversy

In 2021, the FDA approved Aduhelm (aducanumab) for the treatment of dementia from Alzheimer's disease, a devastating disease and the 6th leading cause of death in the US. Its cause is debatable, but irreversible damage to brain neurons occurs, resulting in progressive memory loss and a decline in cognitive function. Before 2021, no drug for the treatment of dementia from Alzheimer's disease had been approved since memantine in 2003.

Often, for controversial drugs, the FDA convenes an advisory committee—a group of medical, scientific, and statistical experts—to review the drug's safety and efficacy. The committee will then vote on whether to approve the drug. While the FDA is not obligated to follow the committee's advice, it does so approximately 80% of the time. In the case of aducanumab, the committee voted unanimously (10 of 11 voted against, with one voting 'uncertain') to not approve the drug. Nevertheless, the FDA granted 'accelerated approval' to aducanumab, and controversy ensued. Three members of the advisory committee quit in protest. The controversy was so large that a congressional investigation ensued [83].

A Phase 2 and two pivotal Phase 3 trials were submitted to the FDA. The Phase 3 data were conflicting. One trial demonstrated a benefit; the other did not. The Phase 2 study demonstrated a benefit at the highest dose (10 mg/kg), which is the FDA-approved dose. Both pivotal studies were stopped early for futility when the pooled analysis showed no benefit. The FDA concluded the data did not conclusively show a benefit but was "suggestive" of benefit [84]. The problem is that scores of previously developed drugs also decreased amyloid beta protein levels but showed no clinical benefit. These drugs were not allowed to be submitted for regulatory approval. So why was this drug allowed to be submitted, and why was it approved?

Ultimately, the FDA approved aducanumab because of the unmet medical need for new drugs to treat Alzheimer's disease. The FDA did not grant the sponsor full approval but rather 'accelerated approval,' a process that allows early access to a drug for serious diseases when its clinical benefit is uncertain. The FDA believed that aducanumab met all the legal requirements for accelerated approval, which require the sponsor to conduct a post-marketing study to demonstrate ultimate clinical benefit. In 2024, Aduhelm was removed from the market because Medicare and private insurers would not reimburse for it.

This situation has happened before. In 2016, the FDA granted accelerated approval to Exondys 51 (eteplirsen), a treatment for Duchenne's muscular dystrophy. In that approval, the study failed its primary endpoint after 1 year of treatment, but there was an increase in dystrophin, a muscle protein. The study size in this case was also very small, with only 12 boys. This approval remains controversial. In 2024, Aduhelm was voluntarily removed from the market by Biogen, which was rumored to be due to low sales and ongoing controversy.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

A few other notable drugs were approved in the early 2020s. The high cost and approval barriers haven't stopped drug companies from pursuing anti-diabetes treatments. Ozempic (semaglutide) is possibly the most exciting drug of the 2020s. Originally developed by Novo Nordisk as a treatment for Type II diabetes in 2012, it entered clinical trials in 2016 and was approved in the US just a year later. Interestingly, Ozempic was developed as an injectable formulation. An oral formulation was later developed, and rather than having two different forms with the same brand-name, it was given the name Rybelsus. But even more interestingly, in those diabetes trials with Ozempic, patients experienced weight loss. After 68 weeks of treatment, patients taking Ozempic had a 12.7 kg (28 lbs.) weight loss compared with the placebo group. This led Novo Nordisk to redevelop Ozempic at a higher dose to treat obesity. The approved Ozempic doses are 0.5, 1, and 2 mg, compared to a recommended dose of 2.4 mg for Wegovy, which was approved in 2021. Sales soon exploded, and competitors followed. Zepbound (tirzepatide) from Eli Lilly & Co. was approved in 2023. Like Wegovy, there is also another version of Zepbound for treating diabetes (Mounjaro). Some predict that sales for these diabetes and obesity drugs can exceed $100 billion by 2030. This class of drugs will likely be the biggest-selling drug class of all time.

For years, the development of drugs to treat Alzheimer's disease was seen as a graveyard where drugs go to die. Namenda (memantine) was approved in 2004 for the treatment of moderate to severe dementia from Alzheimer's disease. After that, no other new drugs were approved for 17 years, until 2021, Aduhelm (aducanumab), a monoclonal antibody targeting soluble and insoluble amyloid beta, was granted accelerated approval, but not without considerable controversy (see The Aduhelm Controversy). Another anti-amyloid monoclonal antibody that reduced plaque formation and slowed cognitive decline in Alzheimer's patients, Leqembi (lecanemab) received accelerated approval in early 2023, but was later changed to full approval based on clinical evidence: 18-month data in 856 Alzheimer's patients showed a 27% decrease in cognitive decline compared to placebo. A third drug, Kisunla (donanemab), was approved in 2024. These drugs represent a milestone for a disease that is so devastating.

The 2000s saw the introduction of biosimilars, which are biologic drugs that are similar to brand-name biologic drugs, with the approval of Omnitrope (somatropin, a drug used to replace growth hormone) in the EU in 2006 and Zarxio (filgrastim, a drug used to stimulate the production of white blood cells in cancer patients) in the US in 2015. It's challenging to get excited about the rise of biosimilars, as they don't represent true innovation. Biosimilars are like generic copies of small-molecule brand-name drugs, but are not considered generic drugs. Their sales don't represent organic growth. Instead, they cannibalize the sales of other brand-name products on the market. It has been argued that biosimilars will be a real cost-saver to consumers because they break the stranglehold a brand-name drug has on the market by introducing competition. However, it appears that biosimilars are often priced only slightly less than their brand-name counterpart.

As of 2022, biosimilar sales are reported to have reached ~$500 billion. There does appear to be a difference in uptake and usage between the EU and the US. The EU was the first to approve a biosimilar, and many more had been approved by the time the first biosimilar was approved in the US. Today, there are far more approved biosimilars in the EU than in the US. Although biosimilars generate sales for a company, they do so at a price; they divert resources that might otherwise be spent on developing new, innovative medicines. Nevertheless, biosimilars are here to stay and are expected to become more commonplace as more biologics come off patent.

In 2025, the first interchangeable biosimilar was approved. This is a significant development in biosimilars because, in the past, you could be prescribed and started on a biosimilar. Still, you could not be switched from a brand-name drug to a biosimilar. Interchangeability is a separate, higher regulatory hurdle to cross. Interchangeable means you can switch from the brand-name drug to the biosimilar and vice versa. Biosimilars are similar but not identical; interchangeability means they are essentially the same. Poherdy was approved in 2025 as an interchangeable form of Perjeta (pertuzumab), a monoclonal antibody used to treat HER2+ breast cancer in combination with other breast cancer drugs like trastuzumab.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Table 5 Examples of Repurposed Drugs			
Drug name	**Original indication**	**Repurposed indication**	**How New Indication was Discovered**
Zidovudine (AZT)	Cancer	HIV/AIDS	In vitro screening of compound libraries
Minoxidil	Hypertension	Hair loss	Discovery of hair growth as an adverse event
Duloxetine	Depression	Stress urinary incontinence	Pharmacological analysis
Ketoconazole	Fungal infection	Cushing syndrome	Pharmacological analysis
Baricitinib	Rheumatoid arthritis	COVID-19	Machine learning
Rituximab	Cancer	Rheumatoid arthritis	Remission of RA symptoms in cancer patients treated with rituxumab
Table modified from [85].			

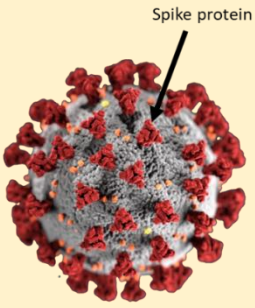
An area that has garnered significant press attention in the 2020s is drug repurposing [85]. There are over 4000 approved drugs in the US. The idea is that some of those approved drugs might be useful for indications not originally studied. This is not like you have a drug approved for use in one cancer type, and now you are trying to see if it might be useful in another cancer type. This is more like you have a cancer drug and now you think it might be useful to treat a viral infection or something completely unrelated to cancer. We've already seen examples of drug repurposing in this book, such as isoniazid as a treatment for depression when it was originally approved to treat tuberculosis. Another notable example is the use of remdesivir, which was initially developed to treat Ebola and later utilized to treat COVID-19. Other examples are presented in Table 5. Historically, drug repurposing was often a happy accident; today, machine learning is frequently touted as a means to identify new uses for existing drugs, and numerous new companies have emerged around this concept.

Some of the challenges around drug repurposing are not necessarily related to finding a drug that can be repurposed, but to the costs associated with proving the drug can be repurposed and then how those costs can be recouped afterwards. Finding a drug that can be repurposed will require clinical trials demonstrating its effectiveness; this may also necessitate dose-ranging studies to determine the optimal dose for a new indication, which are expensive. Furthermore, there is no guarantee that the repurposed drug will be effective in the new indication – the drug may fail in clinical trials. For example, latrepirdine was an antihistamine that Pfizer later studied as a treatment for Huntington's disease. It failed in Phase 3 trials. However, even when the repurposed drug is shown to be efficacious in the new indication, the economics may make it uncommercially viable to conduct these studies in the first place. Often, the patents for these repurposed drugs have expired, and they are now in the generic market. Finding a new indication will require companies to be creative around their patent protection to make the repurposed product commercially viable.

Another notable event was the development of combination therapy for HIV. Recall that in the early 1980s, a diagnosis of HIV was essentially a death sentence, as no drugs were available. In 1987, the first HIV drug was approved, zidovudine (also called AZT), a failed cancer drug from the 1960s. AZT

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

treated the illness but did not cure it, and it could cause liver damage and decrease white blood cell counts, making the individual susceptible to secondary infections. However, by the 1990s, with advances in academic and pharmaceutical research, drugs emerged with different mechanisms of action and varying degrees of efficacy.



Spike Protein on the Coronavirus

Figure reprinted from [86].

The problem with single-agent therapy is that the virus can often mutate, rendering the drugs ineffective. As drug mechanisms evolve, drugs with different mechanisms can be combined to overcome resistance. An early trial in 1995 with AZT and dideoxycytidine (ddC) demonstrated that the combination was more effective than AZT alone [70]. The following year, a study demonstrated that triple combination therapy could reduce viral levels to below the detection limit, thereby preventing the development of mutations and resistance. In 2010, it was shown that healthy patients at risk for HIV could prevent infection by taking antiretroviral therapy as a preventative measure. This combination therapy was called 'highly active antiretroviral therapy (HAART),' and its use decreased the AIDS death rate by 70% since the epidemic's peak in 1995. But with HAART, patients needed to take many pills a day (sometimes as many as 36 a day), and compliance was hard to achieve. Pharmaceutical companies quickly started to develop all-in-one, once-a-day pills containing three drugs to minimize adherence issues. While not perfect, HAART can cause gastrointestinal issues, and resistance can occur if adherence is not 100%. Today, patients with HIV can live full and long lives.

The 2020s also saw the approval of the first non-opioid painkiller. Journavx (suzetrigine) was approved to treat moderate to severe pain in adults and works by blocking sodium channels in nerve fibers, disrupting pain transmission. This approval represents a milestone in pain management that lessens the risk of abuse in patients. 2025 also saw the approval of Yeztugo (lenacapavir), a twice-a-year injection to prevent HIV infection. This may be as close to an HIV vaccine as we will get.

Lastly, I would be remiss if I didn't talk about perhaps the most important drug(s) of the 21st century – the COVID-19 vaccine. By now, everyone knows the story. In late 2019, a small cluster of patients in China developed an unknown, new type of flu that did not respond to conventional treatments. This new flu was later called COVID-19. Despite attempts to limit its global transmission, by March 2020, large parts of the world were under stay-at-home orders. Businesses closed, schools closed, and people began working from home. Children would also attend school from home if possible. By the end of 2020, the US death toll exceeded 200,000. The fate of the world was starting to look grim. Soon after the new flu strain was identified, scientists began searching for a vaccine. Vaccines harness the immune system's power to build an immune response to the antigen. There are lots of ways to do this:

- Live attenuated vaccine contains live virus particles that have been weakened to stop them from causing the disease after administration. Examples include the MMR (measles, mumps, rubella) and chickenpox vaccines.
- Inactivated vaccines contain whole virus particles that have been inactivated or killed to prevent them from causing disease. An example is the polio vaccine.
- Replicating viral vector vaccines use largely harmless, low-pathogenic viruses that are altered to cause the body to produce some of the same proteins as the virus of interest. These are primarily used in veterinary medicine.
- Nonreplicating viral vector vaccines are like replicating viral vector vaccines, except that they cannot replicate in the body and, as such, require higher doses to achieve immunity. An example of this approach is the Ebola vaccine.
- DNA vaccines use DNA plasmids containing the genes for an antigenic protein generated by the disease-causing virus. There are no examples of approved human DNA vaccines.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

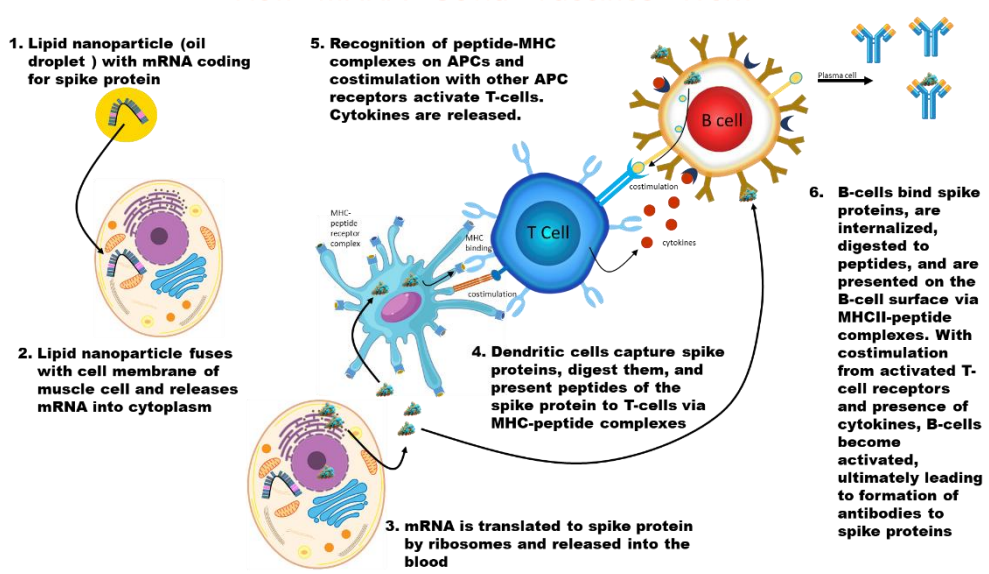


Figure 10. Schematic of how mRNA vaccines work. Elements are not drawn to scale. For example, lipid nanoparticles are significantly smaller than a cell.

- mRNA vaccines use a piece of mRNA that will produce an antigenic protein generated by the antigen of interest to mimic the disease-causing virus. Examples include the Moderna and Pfizer-BioNTech COVID-19 vaccines (Figure 10).
- Subunit vaccines are subunits of the protein of interest that generate an antigenic response independently. They are relatively safe since they do not replicate within the body, but they are not particularly effective and require multiple administrations. Examples include the HPV, pertussis, and hepatitis B vaccines.

Although a live, attenuated vaccine seemed like a logical choice, an mRNA[3] approach was considered to be faster, and speed was paramount given the number of daily COVID-19 deaths. The problem was that no mRNA vaccine had ever been developed or approved for use in humans. Surprisingly, given previous outbreaks of SARS and MERS virus, two other types of coronaviruses, scientists within the US government were preparing for a moment such as this. Building on basic science discoveries developed over the prior decades that have helped understand how HIV works, basic science has shown that mRNA can deliver instructions to cells and synthesize the encoded protein in vivo. Understanding how viruses enter cells, the government and Moderna agreed to develop an mRNA vaccine. The idea was that the mRNA would code for a protein on the capsid of the COVID virus called a spike protein, and that these spike proteins would be sufficiently antigenic to mount an antigenic response to them, such that when the actual virus was encountered inside the body, the body would mount an immune response to the spike protein and kill the virus. Scientists at Pfizer-BioNTech developed their mRNA vaccine almost simultaneously. The race was then on to see who could get approval first. In August 2021, Pfizer-BioNTech received full approval for adults 16 years and older; Moderna received approval shortly after in January 2022. Unfortunately, due to the staggering degree of scientific and numeric illiteracy in the US, many thought the mRNA approach altered their DNA and refused to take the vaccine, leading to an unknown number of deaths that could have been prevented. An excellent account of the development of a COVID-19 vaccine can be found in [87],

[3] Contrary to what some social media influencers state, mRNA vaccines, and mRNA drugs in general, are NOT a kind of gene therapy. They do not interact with your genes or manipulate your genes, or alter your genes in any way. They bypass genes completely in their mode of action.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

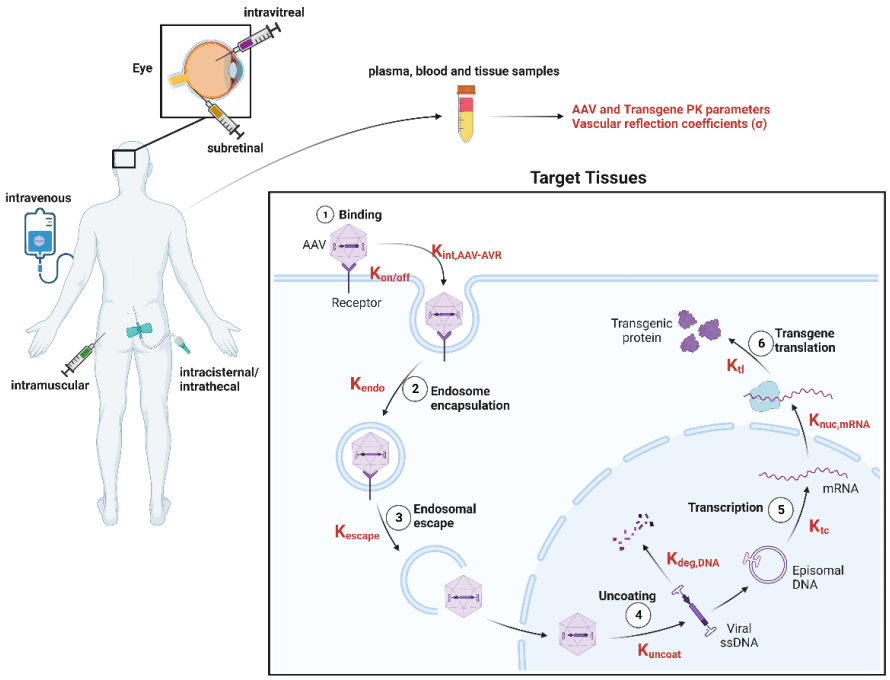


Figure 11. Schematic of gene therapy mechanism of action with various routes of administration. Reprinted from [89] and created using biorender.com.

although a shorter version can be found in [88].

The Future

What does the future of drug discovery look like? That's hard to say, and if I knew, I would be wealthy. It's safe to say that the general movement from small-molecule drugs to molecular biology-based drugs will continue. This is evident from some notable approvals in the last few years. Here are some predictions for new drugs in the decades ahead:

- The Rise of Gene Therapy. To date, several gene therapy drugs have been developed and approved, among them Zolgensma (onasemnogene abeparvovec-xioi), approved in 2019 for the treatment of spinal muscular atrophy in children under 2 years of age. Instead of treating the disease, gene therapy cures the disease with a single injection. Gene therapy wraps a piece of viral DNA inside a protein shell (called a capsid) from another virus, usually an adenovirus (the common cold virus). The capsid merges with the cell membrane and is internalized inside the cell. Once inside the cell, the capsid is uncoated, releasing the viral DNA into the nucleus, where the cell's genetic machinery transcribes it into mRNA, ultimately leading to the creation of the protein of interest. Gene therapy is thought to be most effective against diseases caused by a single protein defect, such as sickle cell disease, which results from a defective hemoglobin protein, or Pompe disease, which is characterized by a defective protein that accumulates glycogen inside the cell, leading to muscle weakness and potentially death. Figure 11 presents a schematic of how gene therapy is thought to work.

  Several problems plague gene therapy. First, the dosing has not yet been determined. The translation of doses that work in animals to humans is not entirely clear, and second, adverse events, like severe liver toxicity, have hampered gene therapy. Because the capsid

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

used to package the viral gene is immunogenic, the body mounts an immune response against it, preventing subsequent doses from having any added benefit. Over the next decade, these problems are likely to be resolved, and by the 2030s, gene therapy is expected to cure many diseases. That leads to the last problem: cost. Gene therapy is expensive, often in the millions of dollars, and there is considerable debate about who will pay for it.

What is the World's Most Expensive Drug?

As of August 2024, the world's most expensive drug is Lenmeldy (atidarsagene autotemcel), which is a gene therapy drug to treat a rare illness in children that is potentially fatal. In 2024, the cost was $4.25 million. This one-time treatment completely cures the illness. I wonder if the parents have to pay a $30 co-pay to their insurance company?

- CRISPR (short for "clustered regularly interspaced short palindromic repeats") allows scientists to selectively modify the DNA of individuals inside their cells in vivo. Taken from the DNA sequences of bacteria, scientists can effectively "edit genes." In 2023, Casgevy (exagamglogene autotemcel) was the first CRISPR-based gene therapy approved by the FDA to treat patients with sickle cell disease. Stem cells are taken from a patient's bone marrow and are edited with Casgevy, creating cells that are normal for hemoglobin. Patients then undergo myeloablative treatment, and the new stem cells are infused back into the patient, leading to the production of new, normal-functioning hemoglobin. The application of CRISPR to new diseases is already ongoing, such as in hereditary angioedema.

  The primary difference between CRISPR and gene therapy is that, with gene therapy, the defective gene remains in the body. With CRISPR, however, the defective gene is effectively removed from the body and replaced with a functioning gene. Some have argued that CRISPR is superior to gene therapy because it is more precise. Certainly, in the decades ahead, this technology may be used to cure many simple, single-point protein diseases – essentially, anything that gene therapy can cure, CRISPR can also do. While certainly exciting and full of promise, CRISPR and gene therapy are limited because there is no evidence that they can be used to cure multifactorial diseases, such as diabetes.

- The COVID pandemic brought the development of mRNA vaccines to the forefront. We can expect to see many more vaccines developed, such as the vaccine for respiratory syncytial virus (RSV). This disease usually causes mild, cold-like symptoms in children but can be life-threatening in infants and the elderly. Arexvy was the first RSV vaccine approved for the elderly. The vaccine uses a subunit of the RSV capsid, similar to many COVID vaccines, prompting the body to mount an immune response.

  The holy grail of vaccines will be the development of an HIV vaccine[4], which has eluded scientists to date because of the rapidity with which the human immunodeficiency virus (HIV) mutates and its ability to hide in the body from the vaccine. I want to say that scientists will soon solve this problem. I do not doubt that at some time in the future, there will be an HIV vaccine, but when is the question? Given the effectiveness of combination drug therapy and the lack of money being spent on HIV vaccine development, it may take a long time for an HIV vaccine to be developed. A COVID-19 vaccine was developed in less than a year because the technology for creating an mRNA vaccine already existed, and the government invested billions in its development. HIV has neither of these things going for

[4] This is what I wrote when I first wrote this chapter. Since then, the Trump administration has decimated the vaccine division at the FDA by firing many scientists involved in the development of new vaccines. I now doubt whether an HIV vaccine will ever be developed.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

it. In the 1990s, when Bill Clinton was President, there was a strong interest in developing an HIV vaccine, but that time has passed. The National Institutes of Health created the Center for HIV/AIDS Vaccine Research (led by Anthony Fauci, the same person who was the chief medical advisor to the President 30 years later during the COVID-19 pandemic), but it has had no success. Nevertheless, research continues along many fronts to find a vaccine; it is just a matter of time before that happens.

Along the lines of mRNA vaccines, mRNA drugs will be developed. For example, instead of administering exogenous recombinant enzyme to patients with lysosomal storage disorders, you could administer the mRNA that encodes the protein, achieving the same effect. Moderna is conducting extensive research in this area.

- Expect a lot more drugs for the treatment of obesity. Bob Ross, the painter, said, "Nothing breeds success like success," and the same is true for drug development. Expect many more companies to enter the field of obesity research, and for additional drugs to follow in the treatment of obesity.
- There might also be a male contraceptive pill. YCT-529, being codeveloped by Columbia University and YourChoice Therapeutics, is the first hormone-free oral male contraceptive and, in 2025, completed Phase 1 trials. In mice, it was shown to be 99% effective in preventing pregnancies, similar to female oral contraceptive pills. Developing a male contraceptive pill has been challenging largely because of the sheer magnitude of what the pill has to overcome. One ejaculation has 300 to 400 million sperm. Even for a drug that is 99% effective, that leaves 3 to 4 million sperm in the ejaculate – and it only takes one sperm to fertilize an egg. Currently, it is 99% effective, which may not be sufficient, but it's on the right track.

Numerous other exciting developments are likely to occur, including the discovery of cures for autoimmune diseases, the development of male contraceptives, the creation of more non-opioid (and potentially non-addictive) painkillers, and numerous drugs for cancer treatment. I can't wait to see what they are!

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 3

Regulatory Considerations

When drugs are developed, companies talk about the "development pathway": the path a drug will take from discovery to approval, including the studies and experiments conducted in between. Alongside this pathway is the regulatory pathway, encompassing all interactions a company has with regulatory agencies throughout the development process. Every country on the planet has regulatory policies in place that must be met before a drug can be sold there. This chapter will discuss general regulatory considerations in the US, with a focus on the FDA. An entire book could be written on regulatory considerations for the three major regions: the US, the EU, and Japan. Adding China to this group might make it the big four, but it is beyond this book's scope to delve into that level of detail. This chapter will also focus on regulatory considerations for new drugs, excluding generic and biosimilar drugs.

The Federal Food, Drug, and Cosmetic Act (FFDCA)

Much has already been written in the chapter on What is a Drug about the FFDCA. Briefly, the original law that established the FDA, the Wiley Act, was largely toothless and focused more on food than on drugs. Following the sulfanilamide disaster, the FFDCA was enacted in 1938, and it remains in effect today. The original law included new provisions that granted the FDA increased regulatory oversight over food, drugs, and cosmetics. Specifically, the law:

- Brought cosmetics and therapeutic devices under the control of the FDA;
- Eliminated the Sherley amendment, which required proof of intent for fraud in misbranding cases;
- Provided standards for the quality of foods;
- Authorized FDA to perform factory inspections; and
- Added court injunctions to previous penalties of seizures or forfeitures.

The FFDCA also required that drugs be safe before marketing access could be granted; evidence of efficacy did not become a requirement until 1962 with the passage of the Kefauver-Harris amendment. Other amendments to the law were made over the years. Here are some examples:

- Regulation of over-the-counter drugs approved in 1972;
- Medical devices were added to the FDA mandates in 1976;
- The Orphan Drug Act for approval of drugs for rare diseases was added in 1983;
- Generic drugs were added in 1984;
- The Food and Drug Administration Act of 1988 established the Department of Health and Human Services (HHS);
- Approval to advertise drugs using legitimate commercial means like television, radio, and print advertisements occurred in 1988;
- Regulation of drugs intended for animals was added in 1988;
- The Prescription Drug User Fee Act in 1992 required companies to pay for review of their

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

products and other services;
- Regulation of supplements was added in 1994;
- The Pediatric Rule, intended to obtain dosing information for children, was added in 1998 and was followed by the Best Practices for Children Act in 2002;
- Project Bioshield in 2004 allowed approval of drugs intended to treat acts of terrorism, like the use of nerve gas agents;
- The FDA Safety and Innovation Act, approved in 2012, was designed to improve the FDA review process.

There have been numerous other amendments since the approval in 1938; the reader is referred to the FDA website for Milestones in Food and Drug Law.

Supporting Players

The interaction between drug companies and the FDA does not operate in a vacuum, with the two main characters just interacting with each other. There is a whole host of supporting players that contribute to the regulatory environment. Here are some of the more important ones.

Institutional Review Boards (IRBs) comprise scientists, non-scientists, community members, and other independent individuals who oversee studies involving human participants. Before a study can begin, an institution must get IRB approval. Multi-center studies generally require approval from multiple IRBs. Their job is to ensure that the study is conducted ethically and in accordance with applicable laws and regulations.

The International Conference on Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) is an independent group comprising regulators and pharmaceutical industry members, whose mission is to develop standards and guidelines for the conduct of clinical research, thereby eliminating the need for companies to meet unique requirements in different countries. Their main guidelines fall into four groups: efficacy (E), safety (S), quality (Q), and multidisciplinary (M). It ensures the alignment of regulatory requirements among participating countries.

Data Safety Monitoring Boards (DSMBs), which the company selects, are independent groups of experts, including scientists, physicians, and statisticians, that ensure the safety of study participants. They monitor a study for safety, efficacy, and data integrity. DSMBs are mandatory for pivotal Phase 3 studies, but are common in other Phase 3 studies and some high-risk Phase 2 studies. DSMBs are usually unblinded to treatment and have the authority to stop a trial for safety reasons or if it is unlikely the drug will show a therapeutic benefit compared to the control arm. They are considered guardians of the trial.

Public Health Service Act for Biological Drugs

While small molecules and some large molecules are regulated under the FFDCA, biologics are regulated under Section 351 of the Public Health Service Act (PHSA) of 1944. Such biological drugs include vaccines, monoclonal antibodies, gene and cell therapies, blood and plasma-derived products, and allergenics. At the time, Congress believed that the complexities of approving a biologic drug necessitated a separate framework from that for small molecules and should not be included in the FFDCA. The original framework for biologic drugs was established in 1902 with the passage of the Biologics Control Act. This act was passed following 13 deaths from tetanus due to contamination of a diphtheria antitoxin by a local manufacturer that lacked quality controls. This act was later folded into the PHSA, which then became the governing law for biologics. One aspect of the PHSA is that the FDA now has the authority to approve not only the biologic but also the manufacturing facility where it is made. This was necessary due to the added complexity in manufacturing biologics compared to small molecules. The PHSA was amended in 2010 to allow for the approval of biosimilars (the 351(k) pathway).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Structure of the FDA

The FDA is responsible for ensuring that food, drugs, and cosmetics are safe, effective, rigorously manufactured, and properly labeled. The FDA also oversees the approval of medical devices and companion diagnostics. In 2024, the FDA regulated more than 100,000 tobacco products, 20,000 prescription drugs, 6,500 medical devices, 1,600 animal products, and 896 licensed biologic products [90]. They also regulate approximately 78% of the food supply, except for meat, poultry, and certain egg products, which are regulated by the US Department of Agriculture (USDA).

The FDA was established with the passage of the Pure Food and Drug Act of 1906, officially known as the Wiley Act. The key provisions of the act were to prevent misbranding or adulteration of food and drug labeling, to ensure drug safety, and to establish the FDA. The act was largely built around food labeling, not around the approval or enforcement of drugs. It wasn't until the 1938 amendment, after the sulfanilamide disaster, that the modern FDA came into existence. It was the two Roosevelts, both progressive Democrats, who signed these laws – first, Teddy and then Franklin D. Roosevelt.

Why is the FDA Called "the Agency?"

Internally and externally, the FDA is referred to as "the Agency." Why? The word 'Agency' is not in its name. It's challenging to pinpoint exactly when the FDA began to be referred to in this manner. It generally began around the early 1960s, coinciding with the Kefauver-Harris amendment to the FFDCA, and it's challenging to pinpoint why it started to be referred to in this manner. Generally speaking, government regulatory bodies are referred to as "Agencies" because they administer statutes passed by Congress. By the 1980s, internal FDA documents began referring to themselves as "the Agency." For example, on the FDA's web page for Background: FDA Good Guidance Practices, it states "*FDA guidance documents explain the agency's interpretation of, or policy on, a regulatory issue." and "FDA uses guidance documents to explain the Agency's current thinking on such matters as the design, manufacturing, and testing of regulated products; scientific issues; content and evaluation of applications for product approvals; and inspection and enforcement policies.*" There is also some broader legal convention at play here to refer to regulatory bodies generically in the third person in the US. Hence, today, you will frequently hear the FDA referred to as "the Agency."

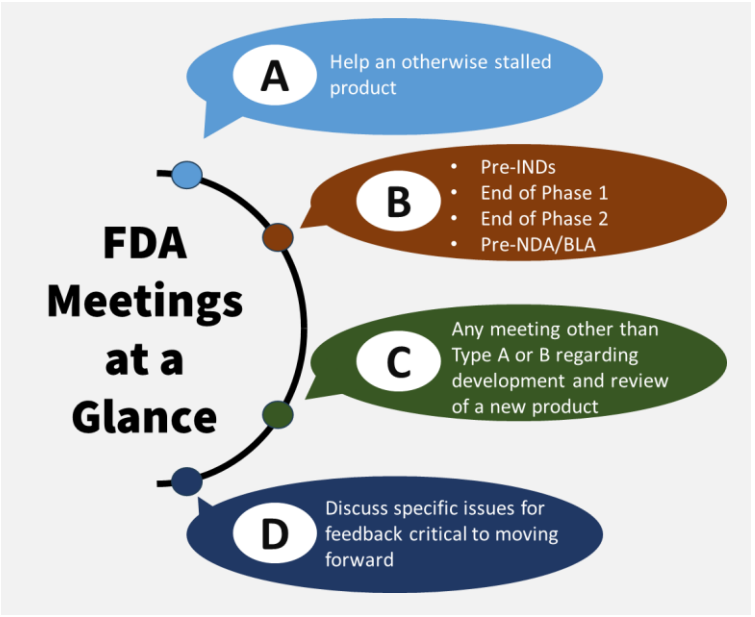
With over 20,000 employees, representing 0.2% of the federal workforce in 2024, the FDA's annual budget was $3.87 billion, accounting for approximately 0.05% of the total federal budget. Regulation of food and drugs accounts for approximately half of the FDA's budget. Using this budget, the FDA regulates 21 cents of every consumer dollar spent in the US, an amount estimated to be more than $3.6 trillion [90]. Today, the FDA operates under the HHS umbrella. The agency, as it is known in the industry, which ironically does not include the word 'agency' in its name, is led by a Commissioner who is appointed by the President and confirmed by the Senate. The Commissioner oversees all the Centers within the FDA.

As of early 2025, before the Trump reorganization, the FDA consisted of nine Centers and 13 offices. It is pointless to print their organization charts here, as they may change; however, the reader is referred to the FDA Organization on the FDA website if interested. Regulation of new drugs is overseen by the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER). Within CDER, there are 13 Offices, like the Office of New Drugs and the Office of Generic Drugs. Within CBER, there are eight offices, such as the Office of Blood Research and Review and the Office of Therapeutic Products. Whether CDER or CBER regulates a product is complicated and is determined by an inter-agency agreement between the two Centers. It depends on their primary mode of action and product class. Generally, CDER is responsible for chemically synthesized molecules, excluding vaccines and allergenics, antibiotics, naturally occurring substances purified from plants, minerals, animals, or solid human tissue sources, certain classes produced by fungi, and hormone products, regardless of their manufacturing method. CBER regulates more complex biologics, including cell therapies, venoms, proteins produced in cell culture, vaccines, blood

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

products, and gene therapies. Either center may transfer review of a product to the other center. That center then becomes responsible for regulating the product and evaluating product quality.

Interactions with the FDA



Companies can interact with the FDA through written correspondence or in-person, video, or teleconference meetings. Notably, the FDA has made substantial efforts over the last 20 years to improve communication with companies. Guidance was issued in 2023 to help sponsors request such meetings [91]. Companies can request advice, seek help with stalled development programs, confirm understanding of guidance, obtain agreement with the FDA on specific development issues such as clinical trial design or the studies supporting the submission, or discuss other issues relevant to the drug. The FDA generally seeks to keep these interactions as efficient as possible due to the sheer volume of work each employee is responsible for. In that regard, the FDA encourages clear and concise questions. There is no guarantee the FDA will approve a meeting request; vague or open-ended questions may result in denial.

After an IND is filed with the FDA, a divisional regulatory project manager (RPM) is assigned to serve as the primary point of contact between the sponsor and the FDA. The RPM also facilitates the timely resolution of questions, conflicts, and communication challenges between the sponsor and the reviewing division. Direct correspondence with reviewers is not recommended. While not subject-matter experts, they will coordinate responses from the appropriate reviewers or the person(s) on the review team.

Generally, there are various types of meetings with the FDA. These are described below:

1. Type A meetings are for drugs that are stalled in development and cannot continue unless they have an FDA agreement. This may include disputes, clinical holds (which will be discussed later), discussions of FDA regulatory actions other than drug approval issues, or discussions of refusal-to-file letters, which are issued when the FDA refuses to review a submission due to incomplete data, rendering it unsuitable for review. Type A meetings are granted to occur within 30 days of the request.
2. Type B meetings can be pre-IND meetings, pre-NDA meetings, certain end-of-phase meetings, such as end of phase 2 meetings, pre-emergency use authorization meetings, or meetings to discuss risk evaluation mitigation strategies. Type B meetings are usually held within 60 days of the request.
3. Type C meetings are any meeting that is not a Type A, B, D, or INTERACT meeting and covers any material not covered by a Type A or B meeting. These meetings are generally considered less urgent and are typically held within 75 days of the request. Examples of Type C meetings may regard pediatric plans, exploratory endpoints or biomarkers, clinical trial design questions outside of pivotal trials, bridging data between populations or formulations, or the use of modeling and simulation to support development programs. Here are some example questions from Type C briefing documents:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Does the Agency agree with the proposed dose-optimization strategy utilizing dose-escalation and dose-expansion from both monotherapy and combination therapy?

Does the Agency agree that Progression Free Survival as an acceptable endpoint for the pivotal trial?

Does the Agency agree with the sponsor's plan to evaluate the effect of renal impairment using clinical data from Study XXXX and using population pharmacokinetic analysis?

4. Type D meetings focus on a few, narrow questions and are limited to involving no more than two FDA reviewing disciplines. An example of a Type D meeting is when the sponsor has a question about the clinical trial design or the statistical analysis of clinical trial data. They might also want further clarification on earlier meetings. For example, here is a sample question from a Type D meeting request:

   The Agency tentatively agreed to the strategy to determine the starting dose in the Pre-IND meeting, pending review of all nonclinical data, but raised concerns regarding the safety profile of the high dose. Therefore, the Sponsor has revised the plan. Does the Agency agree with the revised doses to be evaluated in the first-in-human (FIH) study?

   Type D meetings are designed to provide quick feedback to sponsors, typically held within 50 days of the request. If the questions in a Type D meeting involve more than two reviewing disciplines, it is recommended that the meeting type be changed to Type B or C.

5. INTERACT is short for INitial Targeted Engagement for Regulatory Advice on CBER/CDER ProducTs, and such meetings are intended for early engagement, before the IND or pre-IND meeting. Such meetings are generally focused on IND questions, such as the selection of species for toxicology studies or the use of New Approach Methodologies (NAMs) instead of animal data. Interact meetings started at CBER in 2018 and were only recently accepted by CDER in 2023 under PDUFA VII.

The FDA also has other less frequently used meetings that can be used to communicate with them, such as the Model-Informed Drug Development (MIDD) Paired Program Meeting, which is used to discuss with experts from the FDA Pharmacometrics division model-based approaches to drug development, and Special Protocol Assessment (SPAs) meetings, which are not freely provided by the FDA to the sponsor and require the sponsor to pay user fees to the FDA. SPAs establish an agreement between the sponsor and the FDA regarding the design, endpoints, and statistical analyses for a clinical trial. However, they may also be used for animal carcinogenicity studies, stability protocols for drug products, and animal efficacy protocols under the Animal Rule. Other meetings include CMC-focused meetings and quality insurance meetings related to FDA warning letters, GCPs, and GLPs (BIMO meetings). There are also informal ways to communicate with the FDA. For example, a procedural or clarifying question may be asked of the RPM, such as "Can you confirm if X belongs in both the protocol and the IB?" Such questions may also be quick questions directed towards reviewers. Communication with the RPM is the most common informal channel between the sponsor and the FDA.

Generally, all meeting requests include a briefing book that introduces the product, discusses its background, addresses relevant issues, and then poses questions to the FDA for answers. After the meeting is complete, whether it was a teleconference or face-to-face, the FDA will issue a written set of formal minutes that serve as the official record of the meeting and bind later discussions. The FDA will also issue formal written responses to any query posed to them in a briefing book.

Communication with the FDA is not a one-way street; it flows in both directions. The FDA may issue Information Requests with questions for the sponsor. Examples may include questions about the doses used in a study, details on stopping rules, rules for dose escalation, or other concerns, such as CMC, nonclinical, or data quality issues. Information Requests can come from any part of an IND, NDA, or BLA. Failure to respond to an information request may result in a clinical hold for an IND

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

filing or refusal to approve a drug for NDAs and BLAs.

Types of Applications

Investigational New Drug Applications

In general, by law, all research on new drugs (although there are some niche exceptions [93] in humans, clinical trials must be conducted under an Investigational New Drug (IND) Application, which is the formal request to the FDA to begin clinical trials in humans. The IND contains all preclinical and nonclinical studies, particularly toxicology, to support the starting doses for the first-in-man (FIM) study, as well as all information related to Chemistry, Manufacturing, and Controls around the drug to be administered. The IND also contains the protocol to be followed for the proposed FIM study, along with a copy of the Investigator's Brochure (IB), which summarizes all known information about the drug for investigators in the clinical trial and future studies. The IB is not a static document and must be periodically updated to remain current. A key component of the IND is selecting the doses to be administered in the study.

Companies may meet with the FDA at a pre-IND meeting to discuss any issues before submitting the IND. The FDA has 30 days to review the submission. If the FDA finds no issues with an IND before the 30-day deadline, it will issue a "Safe to Proceed" letter, and the clinical trial may proceed as planned. The FDA does, however, have the right not to issue the Safe to Proceed letter. After the 30 days, the IND is considered approved, and clinical trials may proceed. Federal law prohibits the transportation of an unapproved drug across state lines. IND approval allows the sponsor to ship the drug across state lines.

Components to an IND for Clinical Investigations

1. **Cover letter** (brief explanation of the intended investigation, proposed formulation, etc.)
2. **Forms**: Form 1571 (IND), Form 1572 (Statement of the Investigator), Form 3674 (Certification of Compliance)
3. **Table of contents**
4. **Introductory statement and general investigational plan**: Summary of drug information, previous human experience, and overall clinical development plan (rationale, indication, clinical trial type)
5. **CMC information:**
   - Drug substance: physical, chemical, and biological characteristics, manufacturer, method of preparation, acceptable limits, analytical methods, stability data
   - Drug product: components, manufacturer, manufacturing process, acceptable limits, analytical methods, and stability data)
   - Placebo: composition, manufacture, and control
   - Labeling
   - Environmental analysis requirements
6. **Pharmacology and toxicology Information**: data on pharmacological effect, mechanism of action, drug disposition, toxicology studies (in vitro or in vivo) to support the use in humans
7. **Investigator's brochure**: brief description of drug substance and drug product, summary of relevant preclinical and clinical data, possible risks and side effects
8. **Clinical protocols**: (clinical studies) Form 1572
9. **Previous human experience** with the investigational new drug
10. **Additional information** (if applicable): drug dependence and abuse, radioactive drugs, pediatric studies, and
11. **Relevant information if requested by the FDA**

Reference: [92]

If the FDA has an issue with a study, which can also occur after it initially clears the IND, it may place the study or all studies on clinical hold. Under a complete clinical hold, all studies are immediately stopped; no new patients are enrolled, and no further treatments may be administered to patients. Under a partial clinical hold, parts of a study or entire specific studies in a drug development program are stopped. Examples of clinical hold issues at the IND stage may include disagreements with the

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

choice of starting dose, the highest dose to be studied, the proposed inclusion/exclusion criteria, or CMC issues related to drug purity or stability. After the IND is approved, examples of partial clinical holds include situations in which higher-dose cohorts are prohibited in a dose-escalation study or when a study is stopped. Still, others may continue, such as when adult trials are allowed to proceed but pediatric studies are put on hold. Clinical holds may not be lifted until the clinical hold issues have been resolved.

There are many different types of INDs:

1. A Commercial IND, which is the most common type of IND, is submitted by pharmaceutical companies to study and eventually market an investigational drug.
2. A Research IND is for academic investigators who wish to study an investigational or commercial drug in humans.
3. Investigator INDs are submitted by a physician who both initiates and conducts the study, as well as has immediate supervision over the administration of the drug. This IND might be used for an investigator who wishes to study an approved drug in an unapproved indication.
4. Emergency Use INDs, which were another initiative that arose out of the AIDS crisis of the 1980s, allow a patient or small group of patients with a disease to have access to an investigational drug that might provide therapeutic benefit, and who do not have time to wait for approval of a traditional IND under normal FDA regulations. These are intended for use in emergencies and for life-threatening conditions. An example of a single-use emergency use IND was when Gilead Sciences made remdesivir available to hospitalized patients with severe COVID-19 while it was still in development; the IND was revoked in 2022, when the drug was approved.
5. A Treatment IND is for investigational drugs that look promising in clinical trials for serious or life-threatening diseases, while normal clinical development and approval take place. Treatment INDs enable a broader group of patients to access the drug. For example, a treatment IND might be used for a patient who does not meet the inclusion/exclusion criteria for a clinical trial. An actual example was the use of zidovudine in patients with HIV during the 1980s, allowing patients to be treated before its approval.

More information on INDs can be found on the FDA's website. The FDA has issued guidance on the content and format of INDs [94].

New Drug Applications

Since 1938, New Drug Applications (NDAs) have been the means to which companies can request approval to market and sell a drug in the US. All preclinical and clinical information, including full study reports, integrated analyses across all pivotal trials, and all CMC documentation on how the drug is manufactured, processed, and packaged, is included in the submission. The NDA is designed to provide a complete history of the drug's development. FDA must then determine whether the drug is safe and effective (see the chapter on Efficacy and Safety for what that means in practice).

The NDA must be in a consistent format called the Common Technical Document (CTD), which was approved in 2000 by the International Conference on Harmonisation (ICH) (Figure 12). Many countries, including the US, EU, Japan, Canada, Switzerland, and China, have agreed upon this format. The goal of the CTD was to harmonize submissions across regions, aiding reviews and improving efficiency. ICG Guidelines M4 describe the format of the CTD, while M4Q, M4S, and M4E describe the format of the quality, safety, and efficacy modules of the CTD. To further improve efficiency, data from clinical trials within the CTD generally follow the guidelines of the Clinical Data Interchange Standards Consortium (CDISC), a nonprofit organization that sets standards for data formats. Participation in CDISC is voluntary, but all companies likely follow its recommendations.

To ensure consistency in the review process, the FDA follows its internal criteria, known as the Manual of Policies and Procedures (MaPPs), which are question-based assessments of the drug. For example, in the Clinical Pharmacology Review section of the MaPP, the following questions are asked:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

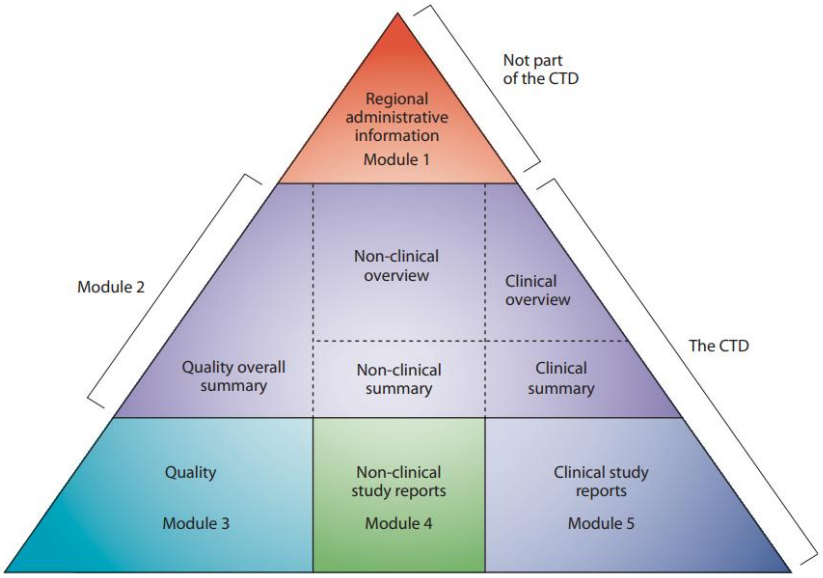


Figure 12. The CTD Triangle. The CTD is the format many countries use for NDAs and BLAs. In 2003, the CTD became required in the EU and Japan, but the FDA strongly recommends it in the US. The CTD consists of 5 modules. Module 1 is region-specific, while the others are to be common across all regions. Figure reprinted from The International Consortium for Harmonisation website.

1. To what extent does the available clinical pharmacology information provide pivotal or supportive evidence of effectiveness?
2. Is the proposed dosing regimen appropriate for the general patient population for which the indication is being sought?
3. Is an alternative dosing regimen and/or management strategy required for subpopulations based on intrinsic factors?
4. Are there clinically relevant food-drug or drug-drug interactions, and what is the appropriate management strategy?

Every reviewer completes their internal questionnaire, and summaries of the reviews are made available after approval at drugs@fda, sorted by drug brand name.

Before the FDA renders its decision on the submission, it may call an Advisory Committee meeting. The FDA has numerous advisory committees that comprise outside, independent experts in the field, including physicians, scientists, statisticians, and lay members of the public, who provide scientific and technical advice on the development and evaluation of drugs. Examples of advisory committees are the Pharmaceutical Science and Clinical Pharmacology Advisory Committee and the Arthritis Advisory Committee. Some committees, such as the Oncology Drug Advisory Committee (ODAC), may recommend whether a drug is safe and effective. Although these committees may make recommendations, their decisions are non-binding; the FDA does not have to follow them, though it often does.

Once the NDA is submitted, the FDA has 60 days to complete a preliminary review and determine whether the "application is sufficiently complete to permit a substantive review." If the FDA determines the application is insufficient, it will issue a Refuse to File (RTF) letter explaining why it did not meet the requirements. RTF letters have generally been confidential correspondence between the FDA and the company. In 2025, to improve transparency, the FDA released over 200 RTF letters for applications that were reviewed between 2020 and 2024. Some of the reasons for rejection include "weak and contradictory evidence", concerns about data generated by third parties, and, in

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

one case, the drug "has an unpleasant odor" that the FDA believed would be problematic for patients being able to take the full dose of the drug, which was an oral solution.

Once the RTF letter period has passed, the FDA must issue a response to the submission in accordance with the Prescription Drug User Fee Act (PDUFA) guidelines. Prior to PDUFA, a review could take 2 to 3 years. PDUFA was started in 1992 and authorized the FDA to collect fees from companies to review their NDA. In exchange, the FDA agreed to review the NDA in a timely manner, typically within 10 months for a standard filing and 6 months for a priority review. These dates are referred to as the PDUFA goal dates. The date the FDA actually issues the final action is referred to as the action date. Typically, the action date precedes the PDUFA goal date, especially for drugs that substantially improve the standard of care.

After the preliminary review is completed, the FDA engages in both formal and informal communication with sponsors. Requests for information may be informally sent to sponsors to clarify review questions throughout the review period. Between Day 74 and 90 for a standard review and between Day 45 and 60 for a priority review, the FDA will send a formal Mid-Cycle Communication with its preliminary review findings. In this letter, the FDA will flag any formal deficiencies or issues that could impact approval and inform the sponsor if any additional analyses or data discussions are required. FDA may also signal whether an advisory committee meeting will be convened. A Late-Cycle meeting will be held around Day 150 for a standard review and between Days 90 and 120 for a priority review, after most of the discipline reviews are complete, to discuss any major deficiencies that could affect approval. Any unresolved safety, efficacy, or CMC questions will be discussed. FDA may also notify the sponsor of the need for an advisory committee meeting, post-marketing requirements, or a Risk Evaluation and Mitigation Strategy (REMS) commitment.

Since new safety information may emerge after the NDA is submitted, sponsors must submit a safety update 120 days after the NDA submission (per 21 CFR 314.50(d)(5)(vi)(b)). The 120-day update applies to both standard and priority reviews. This updated information enables the FDA to have the most comprehensive safety picture of the drug under review. Any substantial change or update during the review, 120-day safety update, or after filing may trigger a 3-month extension to the review time.

Drugs that are not approved receive a Complete Response Letter (CRL) stating that the review is complete and that their application is not approved at this time. The CRL is not a rejection per se, but a refusal to approve in its current form. Sponsors may address any deficiencies, recommendations, or issues and resubmit them at a later date. They may also request a Type A meeting to review the CRL and explore a path forward. Alternatively, the sponsor may withdraw the application without seeking further approval, as technically, a CRL does not terminate the NDA. CRLs are not made public, and it is up to the company to release details.

Biologic License Applications

Biologic License Applications (BLAs) are the NDA equivalents for biologically based drugs, including monoclonal antibodies, cytokines, growth factors, enzymes, immunomodulators, proteins extracted from animals or microorganisms, and other non-vaccine therapeutics. Enzymes or peptides less than 40 amino acids should be filed as an NDA, not a BLA. BLAs generally follow the same format as NDAs and use the CTD format. There are some legal differences between BLAs and NDAs:

- While small molecules are under the authority of the FFDCA, biologics are under the authority of the Public Health Service Act.
- Who reviews a BLA is not entirely straightforward. NDAs are reviewed by CDER at the FDA, while CBER reviews some BLAs. BLAs for monoclonal antibodies, antibody drug conjugates, bispecific antibodies, peptides, and proteins, such as insulin, are reviewed by the CDER.
- NDAs are approved under section 505(b)(1) or 505(b)(2) of the FFDCA, while BLAs are approved under section 351(a).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The FFDCA by the Numbers

If you look up the FFDCA and want to find specific sections of the law, you may find it confusing because of differences in numbering. For example, when the FFDCA was passed in 1938 by Congress, the filing process for new drugs was *internally* numbered as Section 505, lettered subsection b: Filing Application: Contents, and within (b), there are two subsections: (1) for traditional new drugs and (2) for a hybrid pathway. Hence, within the internal numbering system of the FFDCA, drugs may be approved under Section 505(b)(1) or 505(b)(2).

However, if you use the US Code (USC), which is the collection of all permanent laws of the United States, you won't find a Section 505. Within the broad US Code, the FFDCA sits within US Code Title 21 (there are currently 54 titles in total), Chapter 9 (there are currently 29 chapters in total). Specifically, approval pathways are outlined in 21 U.S. Code §355. Hence, under the US Code, 505(b)(1) sits within 21 USC §355(a), and 505(b)(2) sits within 21 USC §355(b)(2). Note that '§' is legal-ese shorthand for 'Section', and §§ refers to multiple sections. You can think of this as follows: people in the pharmaceutical industry refer to Section 505 of the FFDCA, but lawyers refer to 21 USC §355, where Section 505 lives.

The following table maps the FFDCA to the US Code:

FFDCA Section	**Topic**	**U.S. Code Citation**
201	Definitions	21 USC §321
301-310	Prohibited acts and penalties	21 USC §331-337
401-423	Food	21 USC §341-350
501-573	Drugs and devices	21 USC §351-360
505	New Drugs	21 USC §355
601-603	Cosmetics	21 USC §361-364
701-772	General Authority	21 USC §371-379
801-808	Imports and exports	21 USC §381-384
900-920	Tobacco products	21 USC §387
1001-1013	Miscellaneous	21 USC §391-399
351 (Public Health Service Act)	Biologics regulation	42 USC §262
Reference: [95]		

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Supplemental NDAs/BLAs

A supplemental NDA/BLA (sNDA/sBLA) is a post-approval request by the company to change or modify an already existing approved NDA or BLA. Such changes may include modifications to the packaging, manufacturing, dose, formulation, indication, or other aspects of the label. The FDA then reviews the sNDA/sBLA to ensure the change does not compromise the drug's efficacy or safety, or affect the product's quality. When the FDA reviews an sNDA or sBLA, it categorizes it based on the degree of change to the label:

- Major changes, such as any change that might affect the safety or efficacy, e.g., a new indication, dose, or major manufacturing change, require approval by the FDA;
- Moderate changes can be safely implemented while the FDA conducts its review. Of these, there are two types: CBE-0 are Changes Being Effected immediately upon FDA receipt, and CBE-30 are Changes Being Effected 30 days after FDA receipt (unless FDA objects); and
- Minor changes, such as editorial label changes or container size changes, that have minimal impact and can be reported in the next Annual Report.

Urgent safety-related changes to the label may require the sponsor to file an sNDA/sBLA or a CBE-0 to update the label quickly. Timelines for approval of a major change sNDA/sBLA vary, but standard review is within 10 months, or 6 months if priority review is granted.

Abbreviated New Drug Applications

Approved in 1984 through the Hatch-Waxman Amendment to the FFDCA, Abbreviated New Drug Applications (ANDA) use section 505(j) of the FFDCA and are for drugs submitted under the generic pathway for marketing approval. Once approved, the generic drug can be marketed as an alternative to the brand-name drug. These applications are said to be abbreviated because they do not contain any preclinical studies or clinical data from pivotal studies, unlike NDAs and BLAs. Instead, by law, they can reference the original sponsor's studies to support the generic drug's approval. To obtain approval, an ANDA essentially needs to demonstrate bioequivalence with the brand-name drug, typically through a clinical trial, and show appropriate CMC consistent with the brand-name drug, i.e., chemical equivalence. The ANDA pathway was meant to provide low-cost alternatives to brand-name drugs whose patents have expired. The first generic drug for a brand-name drug will receive 6 months of exclusivity before the FDA approves any other generic.

Biosimilar Biologics License Applications

Biosimilar BLAs are the ANDA equivalent for biologic drugs. Unlike other applications, biosimilar BLAs are submitted under section 351(k) of the Public Health Service Act, rather than under the FFDCA. They were designed to create low-cost alternatives for biologic brand-name drugs, which are frequently expensive. Biosimilar BLAs are somewhat different than ANDAs in that the latter require chemical equivalence. Because biosimilars are not exactly the same as the brand-name drug chemically, they must show comparability or "high similarity" with the brand-name drug in regard to CMC. They also require at least one immunogenicity study to assess the biosimilar's immunogenicity and human PK-PD studies to demonstrate what is essentially bioequivalence. However, it is not called bioequivalence, but rather called the "totality of evidence." The first approved biosimilar for a brand-name drug will receive one year of exclusivity before the FDA can approve any other biosimilar.

Drug Approval Pathways

Drugs can be approved through several legal pathways:

1. **FFDCA 505(b)(1):** This is the traditional pathway for new chemical entities, first-in-class drugs, and drugs with novel active ingredients, mechanisms, or indications. The FDA has not previously approved these drugs and requires full reports demonstrating a drug's safety and efficacy. These are stand-alone NDAs. The applicant fully owns the information provided. This pathway is not exclusive to new entities or novel mechanisms. For example, a company may wish to reformulate its product from capsules to tablets. They may choose

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

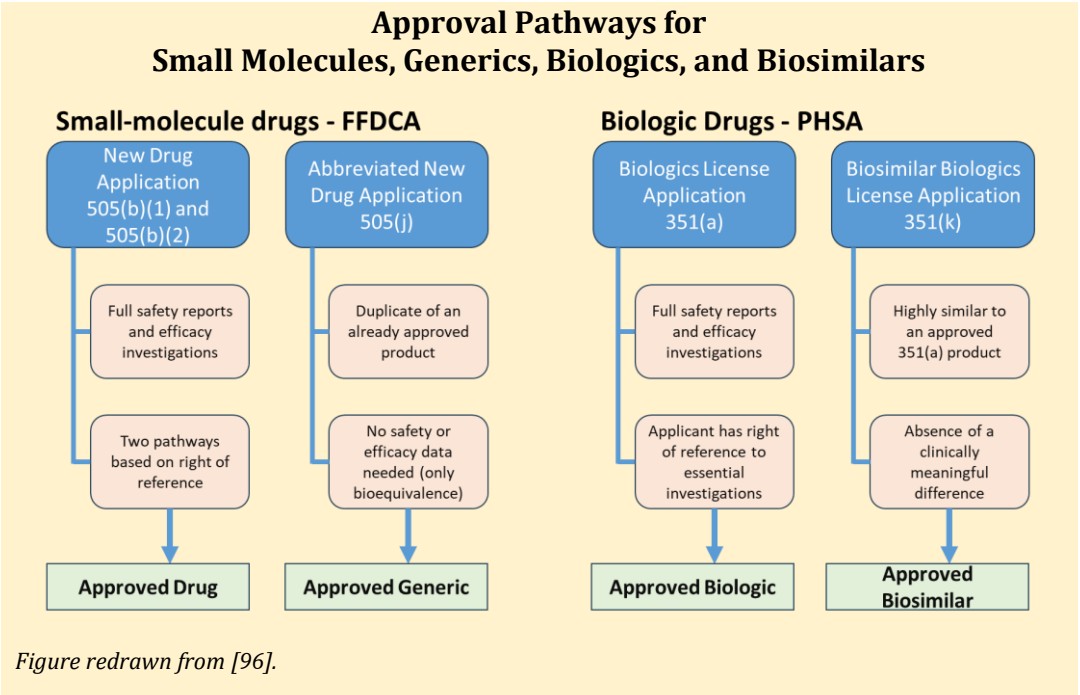


Figure redrawn from [96].

to use the 505(b)(1) pathway to gain approval.

2. **FFDCA 505(b)(2):** This approval pathway is more flexible and could be anything that is not a 505(b)(1) or 505(j), with the sponsor "filling in the gaps" with their own studies, published literature, or prior FDA findings. This pathway can be used for drugs with active ingredients similar to those of a previously approved drug, leveraging information from prior approvals. For example, this may involve a change in formulation, such as switching from an immediate-release tablet to an intramuscular injection, or exploring new uses for existing drugs. This pathway is for drugs closely related to the originator product. The applicant may rely on information already available to the FDA from prior approvals, which the applicant may not have access to, such as full efficacy and safety reports. Legal permission to use data that an NDA sponsor does not own directly and wishes to use is referred to as the Right of Reference and requires a Letter of Authorization from the data owner. An example might be if Company A owns the IND and preclinical package, and Company B wants to reformulate the product. Company B needs Company A's right of reference to the FDA's files.
3. **FFDCA 505(j):** Abbreviated New Drug Applications for approval of generic drugs that can rely on approvals using the 505(b)(1) and 505(b)(2) pathways. The generic drug must have the same strength, dosage form, and route of administration as the brand-name drug.
4. **PHSA 351(a):** Biologics License Applications for approval of biologic drugs.
5. **PHSA 351(k):** Biosimilar License Applications for approval of generic biologic drugs, aka, biosimilars, that rely on using the 351(a) pathway.

There are many ways that the FDA can expedite the regulatory pathway, with maybe one more coming in the future, and these options are often confused with each other:

1. The oldest of the expedited pathways available to companies, **Accelerated Approval** was started in 1992 in response to the AIDS crisis in the 1980s as a discretionary approval pathway available to the FDA, and was made into law in 2012. Accelerated approval allows drugs to be approved based on surrogate endpoints that are reasonably likely to predict clinical benefit, but not necessarily proven to do so. Early uses of this pathway were the approval of drugs used to treat HIV, like AZT (zidovudine), based on the reduction of CD4+

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Expedited Review Options Available to the FDA	
Program	**What it Means**
Fast track	Frequent interaction, rolling NDA submission
Breakthrough therapy	Intensive FDA guidance and all Fast Track benefits
Priority review	Shortened review clock (6 months vs 10 months). All priority reviews are expedited, but not all expedited programs are priority reviews.
Accelerated approval	Approval based on surrogate or intermediate clinical endpoints

lymphocyte counts in HIV patients. Drugs approved by this pathway must perform a confirmatory trial after the drug is on the market. If the confirmatory trial is unsuccessful, the FDA may withdraw the drug from the market. For example, the Phase 3 trial, which tested doxorubicin alone vs. doxorubicin plus the addition of olaratumab failed to demonstrate any clinical benefit. Eli Lilly subsequently removed the drug from the market. Since its inception, approximately 10% of drug approvals have utilized this pathway, with 80% of these being oncology drugs [97].

2. **Priority Review** was granted in 1992 by the Prescription Drug User Fee Act (PDUFA). This act created two types of reviews: standard and priority. The FDA must conduct standard reviews within 10 months of receipt, but priority reviews must be completed within 6 months of receipt. Sponsors may request priority review, but it is up to the FDA to decide whether to grant it. For example, aducanumab for the treatment of Alzheimer's disease was granted priority review in 2020. Generally, priority review is granted if the product provides substantial evidence of greater efficacy or safety than existing approved therapies, or if no current therapies are available. In certain circumstances, such as when a company is developing a treatment for tropical diseases or a countermeasure for biological or chemical weapons, it can receive a voucher for priority reviews. This voucher can be transferred to other drugs or even sold to other companies, sometimes for ridiculous sums. For example, AbbVie paid $350 million to United Therapeutics for a priority review voucher in 2015.
3. **Fast Track**, approved as part of the Food and Drug Modernization Act of 1997, is designed to bring drugs to patients faster by shortening development and review times for drugs used to treat serious conditions that treat an unmet medical need, where no treatment exists or is potentially better than current therapies. Drugs that are granted fast track allow more frequent meetings with FDA to discuss development plans and ensure appropriate data are collected to meet approval requirements, more frequent written communication with FDA, eligibility for priority review and accelerated approval, and rolling review of NDAs or BLAs, which means that rather than waiting until the entire NDA or BLA is received before starting the review, FDA will review sections of the submission as they are submitted in real time. As an example, neflamapimod, used to treat patients with Lewy Body Dementia, was granted fast track in 2019.
4. The newest addition to the expedited regulatory pathways is **Breakthrough Therapy**, which is kind of like fast track but more intensive. Breakthrough therapy requires clinical evidence of superiority over available therapies. In contrast, fast track can be obtained using nonclinical or clinical data and can be granted based on a drug's mechanism of action. The clinical evidence can be based on changes in biomarkers, surrogate endpoints, or a better safety profile with evidence of similar efficacy. A drug that receives breakthrough designation is granted fast-track status, which allows the company greater access to senior FDA staff for development questions and a stronger commitment from senior leaders at the FDA. Although these expedited pathways are generally requested by the company, under

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

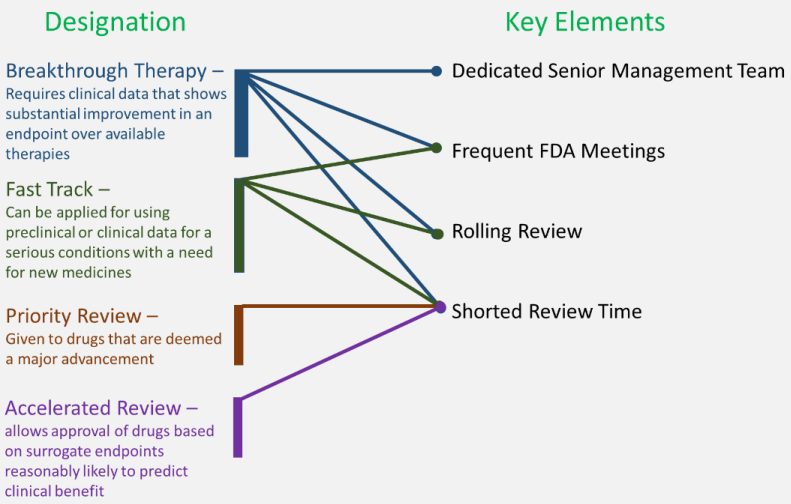


Figure modified from [98].

certain conditions, the FDA may require the sponsor to submit a breakthrough therapy designation request. Ideally, a breakthrough therapy request is submitted no later than the end of Phase 2.

5. A fifth option appears to be available, but its legal status is questionable since Congress did not approve it. In June 2025, the FDA Commissioner announced a new Commissioner's **National Priority Voucher** (NPV) program, under which the FDA will review an NDA or BLA within 1 to 2 months of submission. It is expected that companies will commit to a rolling review of the submission, enabling the final timelines to be established. Drugs accepted into the program are expected to align with national health priorities, such as addressing a US health crisis, meeting unmet public health needs, or delivering more innovative cures to the US population. In October 2025, nine companies were awarded the first NPVs. Of those awarded NPVs, there does not appear to be any underlying theme. However, some analysts have suggested that it may be a reward program for those who align with the administration's interests.

Risk Evaluation and Mitigation Strategy (REMS)

After a drug is approved, the FDA may require sponsors to implement a Risk Evaluation and Mitigation Strategy (REMS). REMS are saved for drugs with serious safety concerns to ensure that the drug's benefits continue to outweigh its risks, and they are not put in place to reduce all adverse events of a drug. Rather, they focus on particular safety aspects of the drug and how to minimize their concerns. They are an additional safety net for high-risk drugs.

For example, the thalidomide disaster is well known and has been discussed often in this book. Despite its notorious history, thalidomide was approved in 1998 to treat leprosy and in 2006 to treat multiple myeloma (Thalomid®). To ensure prevention of birth defects, a REMS program has been put in place (https://www.thalomidrems.com/) using a 3-pronged approach:

- *THALOMID® (thalidomide) is contraindicated in pregnant females and females capable of becoming pregnant. Females of reproductive potential may be treated with THALOMID provided adequate precautions are taken to avoid pregnancy*
- *To avoid embryo-fetal exposure, THALOMID is only available under a restricted distribution program called "THALOMID REMS"*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

- *Only prescribers and pharmacies certified by the THALOMID REMS program can prescribe and dispense THALOMID to patients who are enrolled and meet all the conditions of the THALOMID REMS program*

As per the website, the key points of the program for the prescriber are:

- *The prescriber enrolls and becomes certified with the THALOMID REMS program*
- *The prescriber counsels patient on benefits and risks of THALOMID*
- *The prescriber provides contraception and emergency contraception counseling*
- *The prescriber verifies negative pregnancy test for all female patients of reproductive potential*
- *The prescriber completes a THALOMID Patient-Physician Agreement Form with each patient and sends to the THALOMID REMS*
- *The prescriber/patient completes applicable mandatory confidential survey*
- *The prescriber obtains an authorization number from the THALOMID REMS and writes it on every prescription along with patient risk category*
- *The prescriber writes no more than a 4-week (28-day) supply, with no automatic refills or telephone prescriptions*
- *The prescriber sends THALOMID prescription to certified pharmacy*

And that the pharmacy:

- *The pharmacy certifies with the THALOMID REMS program*
- *The certified pharmacy must obtain a confirmation number from the THALOMID REMS before dispensing*
- *The certified pharmacy counsels the patient, and completes the Education and Counseling Checklist*
- *The certified pharmacy dispenses THALOMID to patient along with a Medication Guide*

There is also a Medication Guide to patients detailing the side effects that can occur with thalidomide. Patients are informed they cannot get pregnant 4 weeks before starting Thalomid, while taking Thalomid, and for at least 4 weeks after stopping Thalomid. Further, patients must agree to use two forms of birth control. They are also informed that these conditions do not apply only to females since thalidomide can pass into semen. Since its inception, no cases of birth defects have been reported.

REMS programs come with black-box warnings on the label, and Thalomid is no exception. The label states:

WARNING: EMBRYO-FETAL TOXICITY AND VENOUS THROMBOEMBOLISM

See full prescribing information for complete boxed warning.

EMBRYO-FETAL TOXICITY

• If THALOMID is taken during pregnancy, it can cause severe birth defects or embryo-fetal death. THALOMID should never be used by females who are pregnant or who could be pregnant while taking the drug. Even a single dose [1 capsule (regardless of strength)] taken by a pregnant woman during her pregnancy can cause severe birth defects.

• Pregnancy must be excluded before the start of treatment. Prevent pregnancy thereafter by the use of two reliable methods of contraception. (5.1, 8.3)

THALOMID is only available through a restricted distribution program, the THALOMID REMS® program (5.2).

FDA maintains a complete list of REMS programs on its website.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Post-Marketing Commitments and Requirements

Post-marketing commitments (PMCs) and Post-Marketing Requirements (PMRs) are commitments the sponsor makes before approval for studies that will be conducted after approval. PMCs are not required by law, but PMRs are. Some programs, such as accelerated approval or the Animal Rule, require a PMR for pivotal studies using endpoints that are not biomarker-based. Although PMCs are conducted voluntarily, the FDA still has enforcement powers if the studies are not conducted as agreed. The FDA establishes PMCs and PMRs to ensure the continued safety, efficacy, and quality of the drug after approval. There are various types, including clinical safety studies, clinical efficacy studies, pediatric studies, pharmacology or nonclinical toxicology studies, and CMC studies. A drug can have multiple PMCs and PMRs at the same time, and companies must report their progress on these to the FDA annually.

In the past, concerns have been raised about companies failing to conduct their PMCs and PMRs in good faith. Since then, the FDA has taken action against companies. The FDAMA of 1997 requires the FDA to publish the performance of a sponsor and their PMCs/PMRs in the Federal Register. If a sponsor fails to meet its PMC requirements, this will damage the company's relationship with the FDA, making future approvals and label negotiations more challenging. The FDA may reject future approvals rather than approve a product with PMCs, as the company has demonstrated it cannot be trusted to meet those commitments. Further, with the FDA posting the status of PMCs and PMRs on its website, companies that fail to meet their commitments may face negative press, potentially damaging their reputations and stock prices. Lastly, if a PMC is critical to ensuring a drug's safety or efficacy, the FDA can convert a PMC to a PMR. If a company cannot meet a PMR, the FDA can issue warning letters, issue financial penalties, or withdraw a drug from the market.

The Prescription Drug User Fee Act

The Prescription Drug User Fee Act, PDUFA for short, was approved in 1992 and is the cornerstone for how the FDA funds and manages drug reviews. The law allows the FDA to collect fees from pharmaceutical and biotechnology companies to review their NDA or BLAs and, in exchange, the FDA agrees to complete the review within certain timelines. In 2025, the application fee for review of a new drug was over $3.5 million. Before PDUFA, the FDA was underfunded and lacked enough reviewers to meet demand, resulting in slow review times. PDUFA injected "User Fees" into the review process, allowing the FDA to hire more reviewers, improve review systems and communication, and set performance goals for faster, more predictable review times.

PDUFA is reauthorized by Congress every 5 years, and new commitments between the FDA and industry, known as PDUFA Goals, are set. The most recent reauthorization is PDUFA VII, which is effective from 2023 to 2027. Some of the key commitments in PDUFA VII include improving use of artificial intelligence tools, data analytics, and cloud-based technologies during the review process, enhancing performance management with better measurement of internal processes and capacity building, large increases in new hires for CDER and CBER, further focus on global regulatory convergence, manufacturing and inspection modernization, and further enhancement of patient-focused drug development using real-world data and digital health technologies, like smart watches.

The impact of PDUFA is undeniable. Review times dropped from years to 6 to 10 months in some cases today. But PDUFA is not without controversy, as some claim it allows industry to have undue influence on the FDA. Critics worry about the pressure on the FDA to approve drugs faster, which may lead to unsafe or ineffective drugs being approved.

Over-the-Counter Drugs

We all buy drugs at our local pharmacy or grocery store: pain killers, drugs for asthma, drugs for allergies, drugs to stop a runny nose, drugs to break a fever, and many other indications. At one point, most of these drugs required a prescription from a physician and had to be filled by a

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

pharmacist. But at some point, they became what are called *over-the-counter* (OTC) drugs. OTC drugs came into existence with the Durham-Humphrey Amendment (1951) to the FFDCA. That amendment distinguished two types of drugs: legend or prescription drugs and OTC drugs that were available without a prescription. Prescription drugs were required to have that statement on the prescription label: "Caution: Federal Law prohibits dispensing without a prescription." A drug does not have to be either-or; a drug can be both a prescription drug and an OTC drug. Often, this difference is due to the formulation's strength. For instance, ibuprofen is available OTC, usually at 200 mg, or by prescription (Motrin, 800 mg). To date, over 300,000 OTC drug products derived from more than 800 active ingredients are available to US consumers.

For a subclass of OTC drugs, these are available as Behind-the-Counter OTCs, meaning you must ask the pharmacist for the medication; they are not strictly available for sale in the store. As a result of the Combat Methamphetamine Epidemic Act (Title VII of the Patriot Act) of 2005, ephedrine, pseudoephedrine, and phenylpropylamine (which is not available in the United States anymore) are more tightly controlled because they can be used as precursors to make methamphetamine. These drugs are still available without a prescription, but they must be sold directly by the pharmacy. There are limits on the amount that can be purchased in a given period and a maximum of 3.6 grams that a person can have on hand at any time. Furthermore, the act requires the purchaser to use a government-issued identification, allowing the state to track the details of the purchase, including who, what, how much, and when, provided the package contains more than 60 mg of the active ingredient. And if you think you can get around these requirements by using a mail-order pharmacy, the limits that can be dispensed at any one time or over a month are less with mail-order pharmacies than with in-store purchases.

Whether a drug can be sold OTC depends on whether consumers can safely use the drug for self-diagnosed conditions, whether the drug does not need a medical professional for safe and effective use, and has a low potential for abuse and misuse. Such changes from a legend drug to an OTC drug are called a Rx-to-OTC switch. Two types of switches from Rx to OTC are possible:

1. Full switch, in which the entire label is changed to an OTC drug in its entirety. Brand-name manufacturers who hold the current prescription drug may petition to an OTC drug by filing an efficacy supplement to an already approved New Drug Application (NDA) or by filing what is called a 505(b)(2) application, which is sometimes called a hybrid application because it allows the applicant to leverage existing clinical data from previously approved drugs or published studies. An example of a full switch is Flonase (fluticasone propionate), which went from prescription to OTC in 2015. The switch to OTC was highly profitable to GlaxoSmithKline; Flonase sales in 2010, before the switch, were around $231 million. After the switch, sales in 2019 exceeded $200 million, indicating minimal revenue loss.
2. Partial switch, in which part of the label is submitted and only some of the conditions of use are made OTC. For this switch, the sponsor must submit an entirely new NDA. An example of a partial OTC switch is the ibuprofen examples noted previously.

An important difference between a prescription and OTC drug is that the latter submission must contain consumer studies demonstrating that consumers can read and understand the label without the help of a medical professional, that consumers can make the right decision on whether it is appropriate to use the medication as directed, that consumers can take the product as intended (don't take that suppository orally, or vice-versa). The submission must also include studies on how consumers use the product in the real world.

There are two ways to obtain an OTC drug designation for drugs that are not already approved as a prescription drug. One way is to file an NDA with the FDA and go through the entire regulatory approval process; this is the hard way. The other route of approval to an OTC drug is by an OTC Monograph. An OTC monograph "*establishes conditions, such as active ingredients, uses (indications), doses, routes of administration, labeling, and testing, under which an OTC drug in a given therapeutic category (e.g., sunscreen, antacid) is generally recognized as safe and effective (GRASE) for its intended use.*" In 1972, OTC Monographs were introduced to assure the public that OTC active ingredients already on the market were safe and effective. This was an enormous effort

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

by the FDA, as they reviewed over 100,000 active ingredients. If a product complies with the rules listed in the monograph, the active ingredient may be used in the product without further approval. Despite the effort put into the monograph system, approximately 20% of active ingredients were never reviewed and remained unfinished over time. In 2020, the Coronavirus Aid, Relief, and Security Act (CARES Act) was approved, updating the monograph system regarding its regulation and approval. It has been estimated that about ¾ of the OTC drugs on the market use the OTC monograph system.

Orphan Drugs

A rare disease is one in which fewer than 200,000 people in the US have it. While this may seem like a small number, a Government Accountability Office report estimates that 30 million Americans have a rare disease, meaning about 1 in 10 Americans has one. That's a staggering number. For decades, drug companies did not venture into rare diseases because it was believed that they could not be developed due to their high development costs; i.e., these drugs would cost too much to develop and not generate sufficient revenue to cover them. In 1983, Congress passed the Orphan Drug Act, which incentivized drug companies to develop drugs for rare diseases. Drugs granted orphan drug status will receive 7 years of marketing exclusivity, meaning no competing drugs for that indication will be approved by the FDA during that period. Furthermore, companies will receive a waiver of their PDUFA fees and can claim tax credits for costs associated with drug development.

The FDA also has a Priority Review Voucher (PRV) Program that entitles sponsors to a priority review (6 months), instead of a standard review (10 months), for a *future* NDA/BLA. PRVs are issued upon approval of the current drug. These PRVs are controversial because they can be sold or transferred to another company, with some sold at upwards of $100 million. Orphan drugs that treat rare pediatric diseases, tropical diseases, or involve medical countermeasures, such as those used to treat public health emergencies, are entitled to such PRVs. An orphan disease can receive orphan status and be eligible for a PRV, but not all orphan drugs are entitled to a PRV.

Four incentives in the Orphan Drug Act encourage companies to develop drugs for rare diseases (orphan drugs):

1. 7 years of patent exclusivity before competitors can enter the market;
2. 25% tax credit for qualified expenses during clinical trials;
3. FDA User Fees waived, up to $2.5 million; and
4. ~$18 million in FDA research grant funding is available

The Orphan Drug Act has been an enormous success. Between 1983 and 1990, approximately 19 orphan drug approvals were made per year; however, by 2020, that number had increased to 34 approvals per year. In 2019, over 500 distinct drugs were approved for orphan indications. Plus, Gleevec showed that approval of a drug to treat a rare disease can still lead to blockbuster profits, so, with the financial incentives for development by the government, it's a win-win for drug companies. And that has been one of the criticisms of the law: companies receive incentives to develop these drugs and then turn around and charge exorbitant prices to patients. Cerezyme has been the poster child for this criticism. Granted orphan drug status in 1994, Cerezyme was approved later that year. The manufacturer, Genzyme, then turned around and charged patients as much as $300,000 per year for treatment [99]. Nevertheless, the law remains, and the situation is what it is. Today, approximately 1 in 5 drugs in development is for rare diseases, benefiting patients.

Access Programs

We've already seen how Emergency Use INDs can be used to provide access to drugs in emergencies and for life-threatening conditions. Patients also have other means of access to investigational drugs.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Expanded Access

Expanded access, also known as compassionate use, is an FDA-regulated program that enables patients with serious or life-threatening conditions who are ineligible for a clinical trial to access investigational drugs. This program requires Informed Consent and Institutional Review Board review and approval. The FDA has three types of expanded access programs: for individuals, for intermediate-sized patient populations, and for widespread use. An example expanded access program is with risdiplam, a drug to treat spinal muscular atrophy, a rare disease that affects motor function and can be fatal. Genentech provided risdiplam to over 150 patients before its approval [100]. The safety data from the expanded access program supplemented the safety database in support of the drug's later approval.

Right to Try

Right to Try (RTT) was approved by the Federal government in 2018. However, many states already had such laws passed by that time that allowed terminally ill patients, who cannot participate in clinical trials and who have considered all other treatments, access to investigational drugs, without FDA involvement. Although it still requires Informed Consent, it does not require IRB oversight. Companies provide the drug directly to patients or physicians. The drug must have started clinical trials, and companies are required to report the results to the FDA. RTT has been criticized as being unnecessary, as other mechanisms are in place to provide drugs to terminally ill patients. Proponents argue that the current FDA mechanisms are too slow and a faster, less bureaucratic method is needed. It's unclear how many patients have used RTT since its approval, but current estimates suggest only a handful have.

Right to Try and Expanded Access are similar but different. Expanded access is granted under the FFDCA (21 CFR 312, Subpart I), whereas the RTT is a stand-alone federal statute (21 U.S.C. § 360bbb-0a). With expanded access, the FDA can modify protocols to put safeguards in place for patients, but with RTT, the FDA is not involved. RTT is meant to be less bureaucratic and faster than expanded access. The FDA receives over a thousand expanded access requests each year, with more than 99% approved for use. In contrast, only about a dozen patients have used the RTT program between 2018 and 2023.

Role of Advocacy Groups in Drug Development and Regulation

The role of advocacy groups in drug development and their impact on regulation cannot be overstated. Starting with the AIDS crisis in the 1980s, advocacy groups changed drug development and how it is regulated. Before the 1980s, it could take years for a drug to get approved. Further, the criteria for approval was strict and onerous. The regulatory pathway was certainly not nimble. People were dying, and the FDA did not seem to care. By 1987, tens of thousands of mostly gay men had died of AIDS, and there were no approved therapies to treat them. One group in particular, Act Up! (AIDS Coalition to Unleash Power), was angry and fed up with the status quo. They would organize and protest at FDA and NIH meetings on the approval process, on the small amount of research being done, and on the lack of drugs available to treat AIDS; often, they were quite verbal and physical, resulting in arrests [101]. Once, they successfully shut down the FDA for an entire day by blocking doorways and roadways, preventing FDA employees from getting to work. They believed that only through civil disobedience could government policies be changed. And they were right. Through their efforts, the FDA altered the way it reviews and approves drugs. The FDA began allowing surrogate markers, such as CD4+ T-cell counts, as measures of efficacy rather than as pivotal endpoints like survival. The approval process was shortened from years to months for drugs to treat life-threatening diseases, and clinical trials became more patient-focused.

Parallel-track, later expanded to include expanded access, allowed patients to access experimental drugs that were shown to be safe, even if efficacy data had not yet been generated. And lastly, advocates demanded a seat at the table – they wanted to be involved in the development and

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

regulation of the drugs that were going to affect them. Today, representatives from advocacy groups serve on numerous FDA committees, providing feedback and offering voting recommendations on drug approvals, as well as assessing FDA policy changes.

Patient advocacy groups also advocated for change before Congress. Without patient groups, laws like the Orphan Drug Act would not have occurred. Greater funding is now budgeted for rare disease research. And drugs like Exonsys 51 or adacanumab would not have been approved without the pressure of advocacy groups. Today, there is a patient-rights movement that patients should be able to determine their own treatment and have access to experimental drugs if need be. Of course, this patient-rights movement does not apply to access to abortion drugs for women, but that is a different story. It was out of this patient rights movement that the concepts of Expanded Access and Right to Try emerged.

Advocacy groups have also influenced how drug development is conducted by pressuring drug companies to align their research priorities with those of patient groups and their needs. Of course, since treatment of rare diseases has been shown to be a lucrative venture, drug companies are none too short of a reason to oblige them. Patient groups have long argued that the endpoints some companies use to assess efficacy, such as progression-free survival, do not capture what matters most to patients, like quality-of-life endpoints, including pain, fatigue, or mobility. Hence, patient-reported outcomes around quality of life are now routinely collected in clinical trials. In return, advocacy groups are a major source of patients for clinical trials, helping companies enroll participants. They might also even help fund research to support drug companies developing a drug that might benefit them. Clinical trials have become patient-centric, making these studies more patient-friendly. For example, decentralized trials allow blood draws and patient visits to occur at a local doctor's office, rather than at a central hospital, where patients would normally go for follow-up. In a sense, patients have become partners in drug development. A very good article discussing the rise of patient-advocacy groups can be found in [102].

The Package Insert, aka the Product Label, aka Prescribing Information

The product label originated with the passage of the FFDCA of 1938, which required that a label contain the same information as the New Drug Application used for drug approval. The label, to be attached to the product or on its packaging, contained information on the proper use of the drug, dosage, route of administration, potential hazards, side effects, and contraindications. Notice it says nothing about efficacy – that didn't come until 30 years later. At this time, the product label was the official name for this information.

With the passage of the Fair Packaging and Labeling Act in 1966, the law requires manufacturers of products that cross state lines to label them honestly and informatively with "full disclosure." This is when those long pieces of paper containing product information were first placed inside the packaging. The Package Insert, as it came to be called, gets its name from the fact that the label was packed into the box, similar to how a prescription medicine is inserted into its packaging. With this new format, the package insert was designed to provide physicians with the necessary information to prescribe the medication properly. It was considered part of the label, even though it was separate from the label attached to the packaging or container. The insert also had to include information on the drug's efficacy.

Over the years, the package insert has become a large, bulky, and onerous document to read. In 2006, the FDA revised the label to make it more user-friendly. The new labels are now referred to as 'prescribing information,' but the old terms' package insert' and 'product label' are used synonymously and can be confusing. For purposes of this book, the terms 'product label' and 'package insert' will be used, and both refer to Prescribing Information. The new format consists of three main parts: highlights of the prescribing section and boxed warnings, the table of contents, and the full prescribing information, which contains all the detailed information about the drug (Figure 13).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

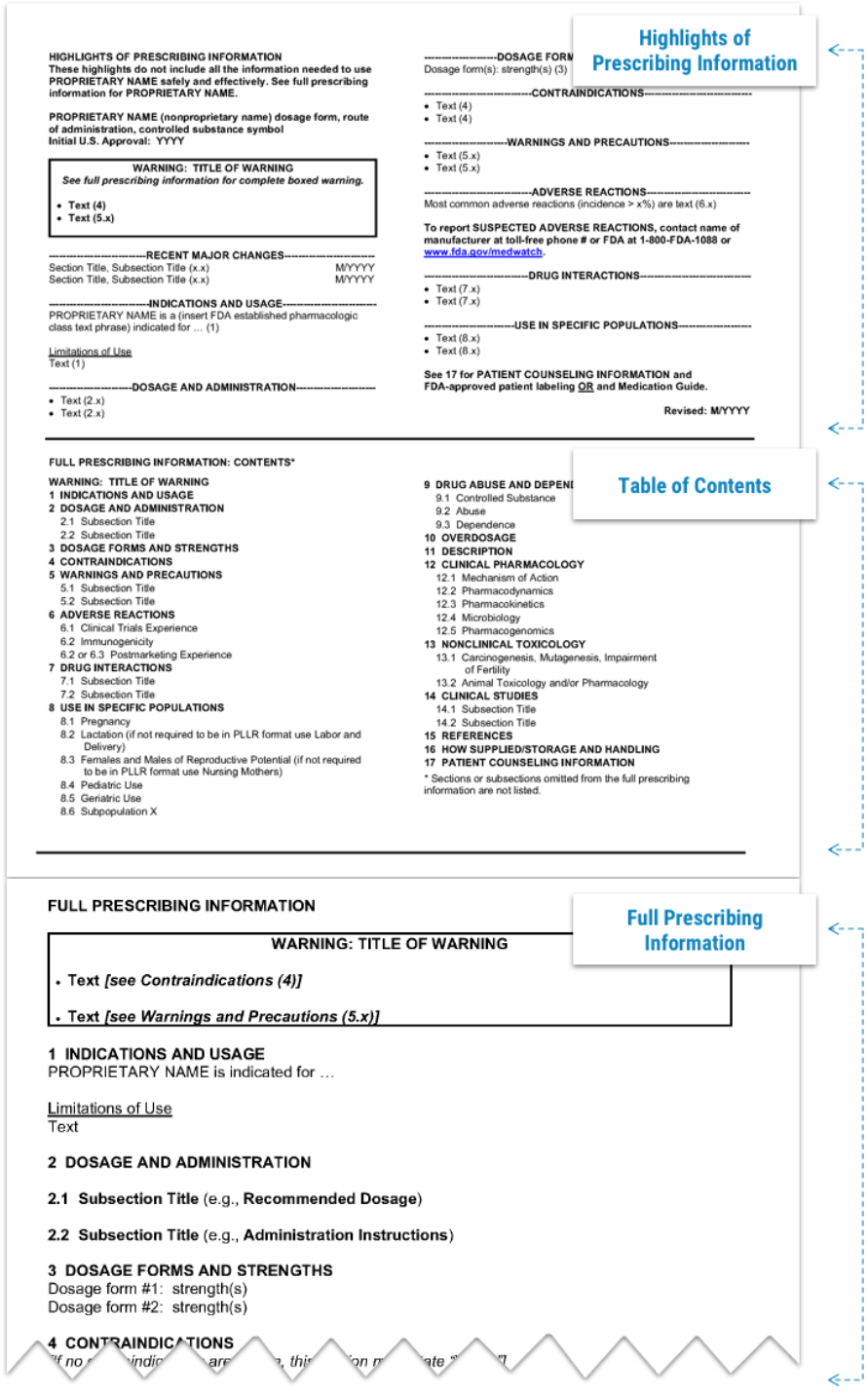
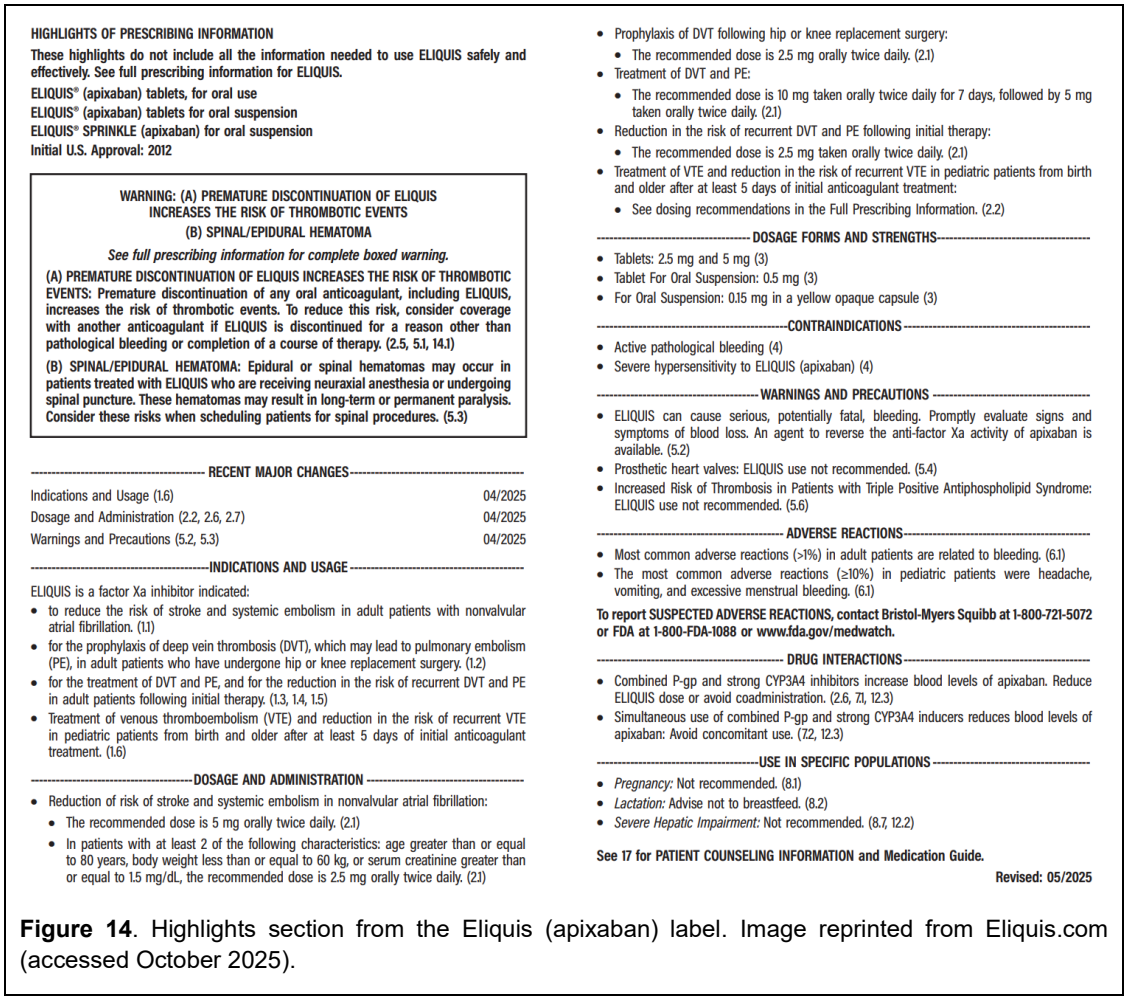


Figure 13. Representative prescribing information under the current format required by the FDA. Figure reprinted from the Food and Drug Administration.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



HIGHLIGHTS OF PRESCRIBING INFORMATION

These highlights do not include all the information needed to use ELIQUIS safely and effectively. See full prescribing information for ELIQUIS.

ELIQUIS® (apixaban) tablets, for oral use
ELIQUIS® (apixaban) tablets for oral suspension
ELIQUIS® SPRINKLE (apixaban) for oral suspension
Initial U.S. Approval: 2012

WARNING: (A) PREMATURE DISCONTINUATION OF ELIQUIS INCREASES THE RISK OF THROMBOTIC EVENTS
(B) SPINAL/EPIDURAL HEMATOMA

See full prescribing information for complete boxed warning.

(A) PREMATURE DISCONTINUATION OF ELIQUIS INCREASES THE RISK OF THROMBOTIC EVENTS: Premature discontinuation of any oral anticoagulant, including ELIQUIS, increases the risk of thrombotic events. To reduce this risk, consider coverage with another anticoagulant if ELIQUIS is discontinued for a reason other than pathological bleeding or completion of a course of therapy. (2.5, 5.1, 14.1)

(B) SPINAL/EPIDURAL HEMATOMA: Epidural or spinal hematomas may occur in patients treated with ELIQUIS who are receiving neuraxial anesthesia or undergoing spinal puncture. These hematomas may result in long-term or permanent paralysis. Consider these risks when scheduling patients for spinal procedures. (5.3)

RECENT MAJOR CHANGES

Indications and Usage (1.6)	04/2025
Dosage and Administration (2.2, 2.6, 2.7)	04/2025
Warnings and Precautions (5.2, 5.3)	04/2025

INDICATIONS AND USAGE

ELIQUIS is a factor Xa inhibitor indicated:

- to reduce the risk of stroke and systemic embolism in adult patients with nonvalvular atrial fibrillation. (1.1)
- for the prophylaxis of deep vein thrombosis (DVT), which may lead to pulmonary embolism (PE), in adult patients who have undergone hip or knee replacement surgery. (1.2)
- for the treatment of DVT and PE, and for the reduction in the risk of recurrent DVT and PE in adult patients following initial therapy. (1.3, 1.4, 1.5)
- Treatment of venous thromboembolism (VTE) and reduction in the risk of recurrent VTE in pediatric patients from birth and older after at least 5 days of initial anticoagulant treatment. (1.6)

DOSAGE AND ADMINISTRATION

- Reduction of risk of stroke and systemic embolism in nonvalvular atrial fibrillation:
  - The recommended dose is 5 mg orally twice daily. (2.1)
  - In patients with at least 2 of the following characteristics: age greater than or equal to 80 years, body weight less than or equal to 60 kg, or serum creatinine greater than or equal to 1.5 mg/dL, the recommended dose is 2.5 mg orally twice daily. (2.1)
- Prophylaxis of DVT following hip or knee replacement surgery:
  - The recommended dose is 2.5 mg orally twice daily. (2.1)
- Treatment of DVT and PE:
  - The recommended dose is 10 mg taken orally twice daily for 7 days, followed by 5 mg taken orally twice daily. (2.1)
- Reduction in the risk of recurrent DVT and PE following initial therapy:
  - The recommended dose is 2.5 mg taken orally twice daily. (2.1)
- Treatment of VTE and reduction in the risk of recurrent VTE in pediatric patients from birth and older after at least 5 days of initial anticoagulant treatment:
  - See dosing recommendations in the Full Prescribing Information. (2.2)

DOSAGE FORMS AND STRENGTHS

- Tablets: 2.5 mg and 5 mg (3)
- Tablet For Oral Suspension: 0.5 mg (3)
- For Oral Suspension: 0.15 mg in a yellow opaque capsule (3)

CONTRAINDICATIONS

- Active pathological bleeding (4)
- Severe hypersensitivity to ELIQUIS (apixaban) (4)

WARNINGS AND PRECAUTIONS

- ELIQUIS can cause serious, potentially fatal, bleeding. Promptly evaluate signs and symptoms of blood loss. An agent to reverse the anti-factor Xa activity of apixaban is available. (5.2)
- Prosthetic heart valves: ELIQUIS use not recommended. (5.4)
- Increased Risk of Thrombosis in Patients with Triple Positive Antiphospholipid Syndrome: ELIQUIS use not recommended. (5.6)

ADVERSE REACTIONS

- Most common adverse reactions (>1%) in adult patients are related to bleeding. (6.1)
- The most common adverse reactions (≥10%) in pediatric patients were headache, vomiting, and excessive menstrual bleeding. (6.1)

To report SUSPECTED ADVERSE REACTIONS, contact Bristol-Myers Squibb at 1-800-721-5072 or FDA at 1-800-FDA-1088 or www.fda.gov/medwatch.

DRUG INTERACTIONS

- Combined P-gp and strong CYP3A4 inhibitors increase blood levels of apixaban. Reduce ELIQUIS dose or avoid coadministration. (2.6, 7.1, 12.3)
- Simultaneous use of combined P-gp and strong CYP3A4 inducers reduces blood levels of apixaban: Avoid concomitant use. (7.2, 12.3)

USE IN SPECIFIC POPULATIONS

- *Pregnancy:* Not recommended. (8.1)
- *Lactation:* Advise not to breastfeed. (8.2)
- *Severe Hepatic Impairment:* Not recommended. (8.7, 12.2)

See 17 for PATIENT COUNSELING INFORMATION and Medication Guide.

Revised: 05/2025

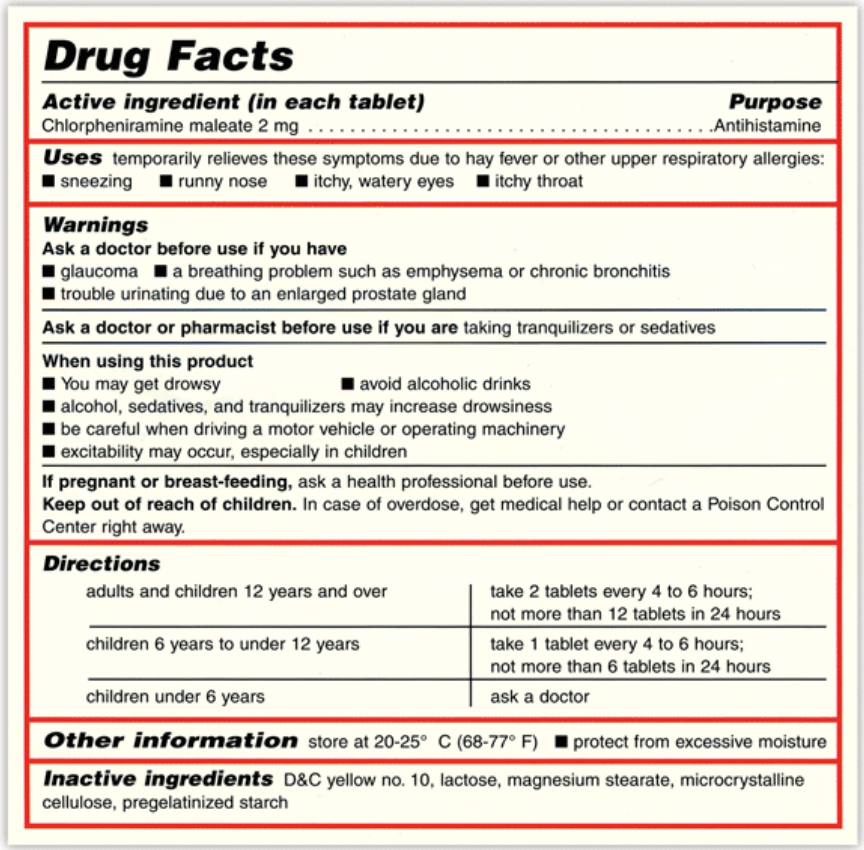
Figure 14. Highlights section from the Eliquis (apixaban) label. Image reprinted from Eliquis.com (accessed October 2025).

The Highlights section is a Cliff Notes version of the Full Prescribing Information (Figure 14). It contains any black box warnings, which are meant to highlight any serious adverse events that may occur with the drug, as well as a concise summary of particular importance from specific sections of the Full Prescribing Information. The Table of Contents is self-explanatory. The Full Prescribing Information features its own format and table of contents, making it easy to find the information you need.

The sections are as follows:

1.) Indications and Usage
2.) Dosage and Administration
3.) Dosage Forms and Strength
4.) Contraindications
5.) Warnings and Precautions
6.) Adverse Reactions
7.) Drug Interactions
8.) Use in Specific Populations
9.) Drug Abuse and Dependence
10.) Overdosage
11.) Description
12.) Clinical Pharmacology
13.) Nonclinical Toxicology
14.) Clinical Studies
15.) References
16.) How Supplied/Storage and Handling
17.) Patient Counseling Information

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Drug Facts

Active ingredient (in each tablet) — **Purpose**
Chlorpheniramine maleate 2 mg ..... Antihistamine

Uses temporarily relieves these symptoms due to hay fever or other upper respiratory allergies:
■ sneezing ■ runny nose ■ itchy, watery eyes ■ itchy throat

Warnings
Ask a doctor before use if you have
■ glaucoma ■ a breathing problem such as emphysema or chronic bronchitis
■ trouble urinating due to an enlarged prostate gland

Ask a doctor or pharmacist before use if you are taking tranquilizers or sedatives

When using this product
■ You may get drowsy ■ avoid alcoholic drinks
■ alcohol, sedatives, and tranquilizers may increase drowsiness
■ be careful when driving a motor vehicle or operating machinery
■ excitability may occur, especially in children

If pregnant or breast-feeding, ask a health professional before use.
Keep out of reach of children. In case of overdose, get medical help or contact a Poison Control Center right away.

Directions

adults and children 12 years and over	take 2 tablets every 4 to 6 hours; not more than 12 tablets in 24 hours
children 6 years to under 12 years	take 1 tablet every 4 to 6 hours; not more than 6 tablets in 24 hours
children under 6 years	ask a doctor

Other information store at 20-25° C (68-77° F) ■ protect from excessive moisture

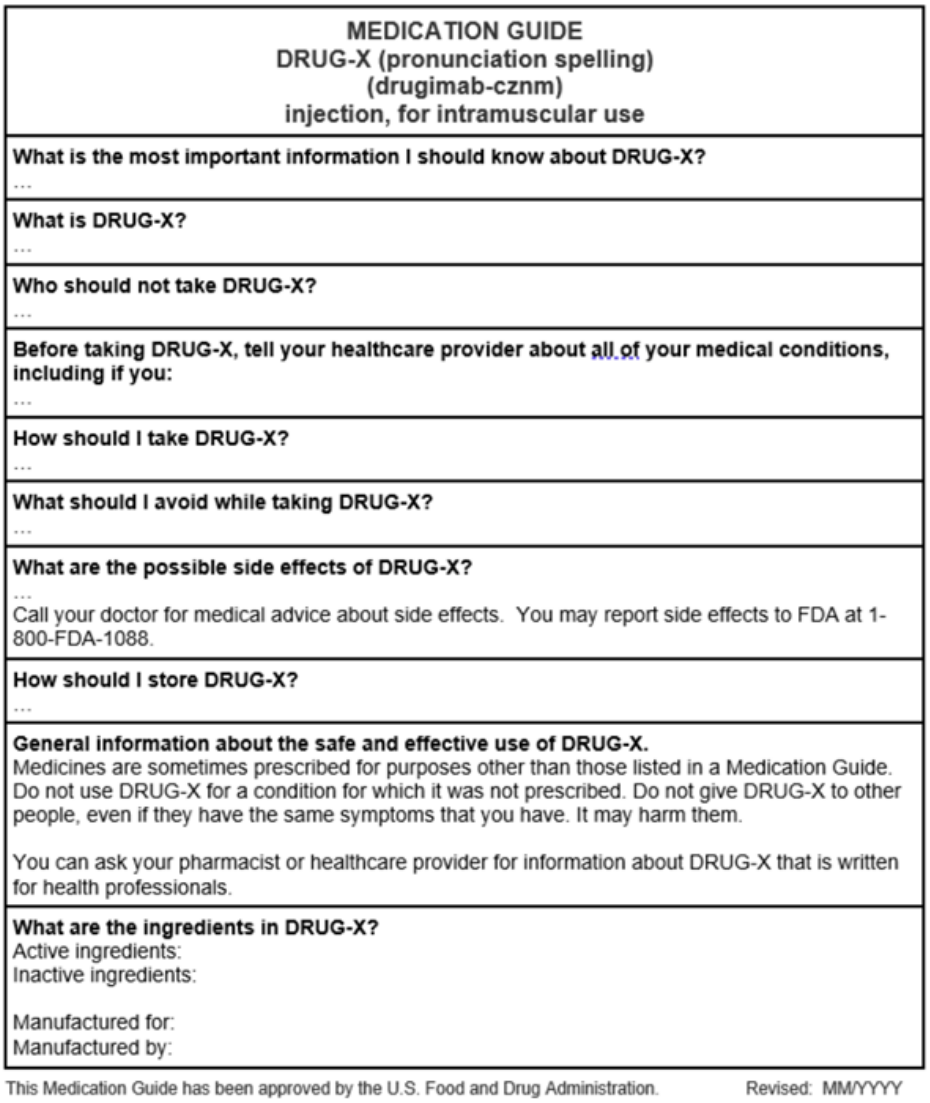
Inactive ingredients D&C yellow no. 10, lactose, magnesium stearate, microcrystalline cellulose, pregelatinized starch

Figure 15: Example drug facts label for an OTC product. Figure reprinted from the Food and Drug Administration.

The actual writing of the insert is the manufacturer's responsibility. The FDA reviews and approves the labels, provides edits, and assures the factual accuracy of the information provided. It is not meant to be promotional. While the FDA has a responsibility to ensure the label meets the requirements of the law, it bears no legal responsibility for the label's content. Interestingly, the legal duty of the label is not to provide information to the patient, but to the prescribing physician. This is based on the learned intermediate doctrine, which assumes that doctors are better equipped to assess the risks and benefits of drugs than patients, and that patients rely on their doctors for information about drugs. If a patient experiences an adverse event while taking a medication, the package insert does not absolve a manufacturer or a physician of legal liability. Both can be sued for malpractice, negligence, or misuse.

This is where it gets murky: physicians are not legally bound to follow the directions on a drug's package insert. A physician can prescribe a medication for an illness or adjust the dose or dosing regimen not found on the label. This is called off-label use, and companies cannot promote any off-label use of the drug [103]. Furthermore, if a drug is approved for use in a particular illness, there is no legal requirement for a physician to prescribe it to a specific patient with the illness. The package insert provides information; it is not legally binding. One of this book's goals is to help the reader better understand the package insert, learn about the information it contains, and how that information was collected.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



MEDICATION GUIDE
DRUG-X (pronunciation spelling)
(drugimab-cznm)
injection, for intramuscular use

What is the most important information I should know about DRUG-X?
...

What is DRUG-X?
...

Who should not take DRUG-X?
...

Before taking DRUG-X, tell your healthcare provider about all of your medical conditions, including if you:
...

How should I take DRUG-X?
...

What should I avoid while taking DRUG-X?
...

What are the possible side effects of DRUG-X?
...
Call your doctor for medical advice about side effects. You may report side effects to FDA at 1-800-FDA-1088.

How should I store DRUG-X?
...

General information about the safe and effective use of DRUG-X.
Medicines are sometimes prescribed for purposes other than those listed in a Medication Guide. Do not use DRUG-X for a condition for which it was not prescribed. Do not give DRUG-X to other people, even if they have the same symptoms that you have. It may harm them.

You can ask your pharmacist or healthcare provider for information about DRUG-X that is written for health professionals.

What are the ingredients in DRUG-X?
Active ingredients:
Inactive ingredients:

Manufactured for:
Manufactured by:

This Medication Guide has been approved by the U.S. Food and Drug Administration. Revised: MM/YYYY

Figure 16: Sample template for a Medication Guide. Figure reprinted from the Food and Drug Administration.

Today, Prescribing Information for physician-prescribed drugs can be found on the website for a specific drug, often at the top of the page, or on Drugs@FDA: FDA Approved Drugs. Although it is generally agreed that the new format improves readability, few studies back this statement. Over-the-counter drugs, which do not require a prescription, have their own label called Drug Facts (Figure 15). In the 1990s, labeling for patients became available, and it comes in different formats: Medication Guides, Patient Package Inserts, Instructions for Use, and Patient Medication Information.

Medication Guides (Figure 16) are used when the FDA determines that patient labeling could improve the product's safety and prevent serious adverse events by directly informing patients of the risks associated with taking the drug. Like the Package Insert, the format of the Medication Guide is prespecified. Patient Package Inserts (PPIs) are required for estrogen-containing products and oral contraceptives. PPIs have their own format and are similar to Medication Guides. PPIs may be voluntarily developed by the manufacturer for other prescription drugs and are not mandated.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Instructions for Use are provided for drugs with complicated dosing regimens or detailed patient-use instructions. These are meant to be step-by-step instructions for the use of the drug. In contrast to other patient-oriented materials, Instructions for Use are not subject to FDA approval. An example is the Instructions for Use provided with the EpiPen. EpiPens are autoinjector syringes that administer a fixed dose of epinephrine to prevent anaphylaxis. The Instructions for Use provide photographs, drawings, and instructions for how to hold the injector, where to administer it, and how to ensure your EpiPen is ready for use.

The latest patient information is the Patient Medication Information guides, proposed in 2023 but not yet officially approved. PMIs are meant for prescription drugs, blood, and blood components that are used, dispensed, or administered in an outpatient setting (like at home). PMIs provide essential information that patients should know about their drug, including how it should be used and administered.

Off-Label Use

General Off-Label Use

When a drug is approved for use by the FDA, it is approved for a specific indication. Sometimes, the indication is quite narrow, in which the drug is said to have a 'narrow label' or 'narrow indication'; alternatively, it can be broad, covering a large population of patients. Companion diagnostics by definition, narrow the label's indication. Physicians are not bound to prescribing a drug only for the approved indication on the label. Physicians can prescribe a drug for any indication they see fit. These unapproved uses are referred to as 'off-label' usage, and it is quite a common medical practice. Some estimates put the figure at roughly 1 in 5 prescriptions in a general doctor's office for off-label use.

The Bottom Line – All patients need to be aware of the drugs they are taking and whether they are being prescribed off-label. If a drug is prescribed off-label, discuss with your physician the rationale for using it off-label and the associated risks.

Why might a doctor prescribe a drug off-label? There are many examples:

- They may have tried every approved drug to treat a patient, and none of them have worked. They may want to try something different to see if that drug would be useful.
- A drug may be approved for a specific cancer, and the physician may want to try the drug in a different cancer type.
- A drug may only be approved for use in adults, but the patient is a child, and the drug may be a useful treatment.
- There is a side-effect of the medication that might be useful to patients. For example, trazodone is used to treat depression, but as a side effect, it makes people tired and sleepy. Sometimes doctors prescribe trazodone for people who have trouble sleeping.
- The drug might be useful in untested indications. For example, β-blockers are used to treat high blood pressure, but by lowering blood pressure, they are also useful for treating social anxiety or fear of public speaking.
- Sometimes, off-label use may not refer to the indication but to something else on the label. For instance, the drug may be approved for a 30 mg dose once daily, and the doctor wants the patient to take 60 mg once daily. Or maybe they want the patient to take the drug twice daily, instead of once a day.

In these off-label uses, it is important to remember that the FDA has not approved their use.

There are several issues related to off-label use that pertain to the patient and the doctor. First, if an adverse event occurs, the doctor may not necessarily be held liable for negligence or malpractice. The doctor may be liable if they failed to adequately disclose the risks, resulting in a lack of informed consent, or if the drug did not meet the standard of care, a somewhat nebulous term. Some courts have held that a doctor may be liable for malpractice if a reasonable patient would have declined

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The United States Pharmacopeia

The US Pharmacopeia (USP) is one of those supporting players to the FDA. They are an independent, nonprofit organization that is over 200 years old. Its mission is to improve public health by developing and disseminating public standards for quality, safety, and benefit of medicines (and other health products/ingredients). Every 2 months, the US Pharmacopeia updates its pharmacopeia, also known as the USP (which is confusing), and the National Formulary (USP-NF). The USP-NF contains a list of over 5000 drugs in the form of a monograph that have USP quality standards. These standards define the identity, strength, quality, and purity of medicines. Drugs subject to USP standards include human and animal drugs. The NF contains monographs for excipients. The USP-NF is recognized by law under the FFDCA as an "official compendia" and is used by more than 140 countries outside the US. Drugs sold in the US must conform to quality standards as defined by the USP-NF. Drugs listed in the USP-NF are considered adulterated if their purity, quality, or strength does not meet the standards defined in the monograph, even for the original manufacturers of the drug. USP has no role in enforcement; that is the FDA's role.

treatment had the risks been adequately disclosed beforehand. They cannot, however, be held liable for nondisclosure. Although many ethics groups urge disclosure, doctors have no legal duty to tell a patient the drug they are prescribing is off-label. Furthermore, some drugs prescribed off-label may not be covered by insurance companies, and the patient may be responsible for the cost.

Until 1997, a company was strictly prohibited from promoting any off-label use. That changed with the passage of the Food and Drug Modernization Act of 1997 and with later legal challenges. Today, off-label promotion is complicated. Generally, although a doctor can prescribe a drug off-label, companies are not able to directly advertise a drug's off-label use. Still, they can discuss off-label usage with a physician if the physician first solicits the discussion. Furthermore, companies can distribute copies of peer-reviewed articles and book chapters on the use of drugs beyond their approved indications to support continuing medical education. Off-label promotion is considered misbranding by the FDA and in violation of the FFDCA; it is one of the major reasons the FDA fines companies. For example, in 2012, Abbott paid $1.6 billion in penalties for off-label marketing of valproic acid. That same year, GlaxoSmithKline paid $3 billion in penalties for off-label marketing of paroxetine for use in children, despite the indication being approved only for adults.

You may be asking, Why doesn't a company apply to have the off-label use on the label? Often it comes down to cost. The cost to do the clinical studies necessary to demonstrate safety and efficacy for the off-label indication is frequently prohibitive. Sometimes, the off-label indication emerges after the drug's patent has expired. There is then absolutely no incentive to conduct the necessary studies for approval. An example of this is the use of doxycycline to treat Lyme disease. There is no label indication for this, but doxycycline is not standard of care for patients with Lyme disease. Sometimes off-label use does not align with the company's product strategy. For example, the sleepiness side effect seen with trazodone is not beneficial for patients, so advertising it on the label would be detrimental to the product overall. Nevertheless, trazodone is prescribed by physicians to treat insomnia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 6 Important Distinctions between 503(A) and 503(b) Compounding Pharmacies		
Distinctions	**503(A) Facility – Local Compounding Pharmacy**	**503(b) Facility – Outsourcing Pharmacy**
Stipulations	Must have a valid prescription for an individual patient	Can produce batches of compounded drugs for future distribution
Distribution	Can be distributed to other pharmacies controlled or owned by that pharmacy within a 1-mile radius	No shipping restrictions
Limitations	Maximum 30-day supply for one patient	None really
Regulating body	State board of pharmacy	FDA
Must register with the FDA	No	Yes
Advantages	Can be patient-tailored	Cannot be patient-tailored, but is cost-efficient; bulk supply available
Compliance with GMP	No	Yes
Production of a drug that is essentially a copy of a drug	Not permitted	Not permitted
Table modified from [104].		

Compounding

An important off-label use that needs to be discussed is compounding, which is when a pharmacist, physician, or someone supervised by a pharmacist or physician creates a formulation for an approved drug that is not commercially available, such as creating a topical cream for a drug normally sold only as a tablet, or when the formulation is made directly from the drug product, such as mixing a drug directly into an intravenous bag. Compounding dates back to the earliest times, reflecting how pharmacists prepared drugs directly for patients. Compounding still requires a physician's prescription, and physicians may be held legally liable if the use of a compounded product results in an adverse event, if the formulation is not pure, sterile, and free of contaminants, or if it contains the wrong dose. Still, compounding may be of value when commercial products do not meet a patient's needs. Compounding became a significant issue in 2025 due to a manufacturing shortage of GLP-1 agonists for weight loss. Some companies stepped in and compounded the drugs directly for patients.

Compounded drugs are not FDA-approved; they are considered off-label. They do not have the same quality standards that approved drugs have. Sometimes, the drug being compounded can be acquired by the manufacturer of the approved drug; other times, they cannot. When they are not active, drug ingredients are purchased by manufacturing companies "approved" by the FDA. These companies can be located anywhere in the world, but they are most commonly in the US, China, and India. Often, this approval involves simply filing a form stating the product will be manufactured there.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 7

Definitions and Examples of the Five Types of Drug Schedules in the US

Type	Descriptions	Examples
I	High potential for abuse No currently acceptable medical use Lack of accepted safety for the use of the drug under medical supervision	Heroin LSD Marijuana
II	High potential for abuse Currently accepted medical use Abuse may lead to severe physical or psychological dependence	Methadone Morphine Cocaine
III	Less potential for abuse than Schedule I or II drugs Currently accepted medical use Low to moderate potential for physical and psychological dependence	Nonamphetamine stimulants Some antidepressants Anabolic steroids
IV	Low potential for abuse Low potential for dependence	Diazepam Chlordiazepoxide Meprobamate Phenobarbital
V	Less potential for abuse than Schedule IV drugs Currently accepted medical use Abuse may lead to limited psychological or physical dependence relative to Schedule III drugs Generally used for antidiarrheal, antitussive, and analgesic purposes	Compounds with small amounts of controlled substances Codeine

Compounded drugs are made by compounding pharmacies, which your typical Walgreens cannot do. There have been instances where compounded drugs are a safety concern because of contamination. In 2012, there was an outbreak of fungal meningitis that led to 60 deaths that was traced back to a compounding pharmacy in Massachusetts.

As a result of these deaths, Congress passed the Drug Quality and Security Act in 2013. There are two parts to the law: Part 1 pertains to compounding, while Part 2 addresses drug security and the prevention of fraudulent products. Part 1, the Compound Quality Act, updates provisions in Section 503A of the FFDCA and creates a new section, Section 503B, that relates to facility compounding manufacturers. 503A still concerns mom-and-pop compounding pharmacies meeting individual patient prescriptions. 503A is regulated at the state, not the federal, level. 503B concerns bulk compounding sites, which must follow Good Manufacturing Practices and are under direct FDA oversight. 503A drugs cannot be "essentially a copy" of the approved drug, unless there is a shortage of the drug. For example, a compounding pharmacy cannot create a 10 mg capsule of a drug if a 10

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

mg capsule of the same drug already exists. The FDA publishes a list of drug shortages in the US. If a compound is on the shortage list, it is allowable to compound the drug as "essentially a copy" of the approved drug. Table 6 presents some important distinctions between 503(A) and 503(b) compounding pharmacies.

Drugs of Abuse

When people hear the word 'drug,' most often they think of drugs of abuse, which are those drugs that can lead to physical or psychological dependence, can be abused, or pose a significant health risk to an individual or community. They don't necessarily have to be drugs of medicinal value. Many of the most notorious drugs of abuse, like alcohol, PCP, or heroin, have no medical value. But many drugs of abuse do have medical value. Cocaine, methamphetamine, heroin, and oxycodone (OxyContin) have all been used medicinally at one time or another (though they may not be in use today, e.g., heroin).

Drug Dependence and Addiction

Physical dependence is a state of adaptation from repeated administration and can lead to withdrawal symptoms upon cessation.

Psychological dependence is a need or "craving" for the drug such that users feel that they cannot function without the drug.

Drug addiction is different from physical dependence. Dependence is a physiological response to a drug.

In contrast, ***addiction*** is a behavioral term for when drugs are used in a manner or amount different than acceptable medical use and often results in compulsive, drug-seeking behavior by the addict. Although many addicts are psychologically and physically dependent on the drug they are abusing, physical dependence can occur in the absence of abuse. An example of the latter is the use of narcotics for chronic back pain.

Drugs of abuse are regulated by the Drug Enforcement Administration, which places controls over how they can be dispensed. The Controlled Substances Act of 1970 places all drugs somewhere on a schedule (Table 7). Placement determines whether the drug is controlled or not controlled. Both the FDA and the Drug Enforcement Administration (DEA) are responsible for determining whether a drug is scheduled and where on the schedule it should be placed. Scheduled drugs are placed into one of five groups, with Schedule I drugs having the most severe dependence and no accepted medical use. Schedule II to V drugs all have an accepted medical use, but have different levels of physical and psychological dependence. Illegal dispensing or possession of scheduled drugs may result in fines and imprisonment.

How Do Drugs Get Their Names?

Drugs have many names. All small chemicals have their chemical names defined by the International Union of Pure and Applied Chemistry (IUPAC). These names are based on the chemical's structure. For instance, acetaminophen has the chemical name N-4-hydroxyphenyl-acetamide. Aspirin is 2-acetylbenzoic acid. Omeprazole's chemical name is 5-methoxy-2-[(4-methoxy-3,5-dimethylpyridin-2-yl)methyl-sulfinyl]-1*H*-benzimidazole. These names are often too lengthy for routine use, so to protect their potential patents and intellectual property, companies use a number or alphanumeric combination instead, e.g., ASP8232 or LY2605541. The letter(s) at the front of the name often denote an abbreviation for the company that developed them, i.e., ASP stands for Astellas Pharma, and LY stands for Eli Lilly & Co.

At such a point, generally at the time of filing of the Investigational New Drug (IND) application, that the company deems the new chemical entity may be a viable drug candidate, they may file for a non-proprietary name with the United States Adopted Names (USAN) Council, a nonprofit organization consisting of the FDA, the American Medical Association, the American Pharmacists Association, and the US Pharmacopeia. USAN works with the International Non-proprietary Name (INN) Committee of the World Health Organization and other groups to standardize the given name. These names are non-proprietary because, once they come off patent, they are free for anyone to use, such as generic

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

drug manufacturers.

USAN names cannot conflict with already registered proprietary and non-proprietary names. There are other rules for valid USAN names, such as the name should be easily recalled, easily spoken, and recognizable, preferably with no more than four syllables; the name should be consistent with other members of the same drug class and have the same stem; and a single word may modify names to account for the salt or ester form of a drug, e.g., cortisone acetate. There are also rules governing the naming of a drug. For example, non-proprietary names cannot imply a drug is "better," "newer," or "more effective." The letters h, j, k, and w cannot be used because they lead to pronunciation problems in some countries. Prefixes cannot imply a chemical or compound, such as calcium. There are too many rules to list here, but it is clear that the naming process is complex, and the USAN frequently rejects names proposed by companies. In such a case, USAN will suggest an alternative name; however, the sponsor also has the right to suggest other names. However, this restarts the review process or initiates arbitration. In rare cases, a settlement requires court review.

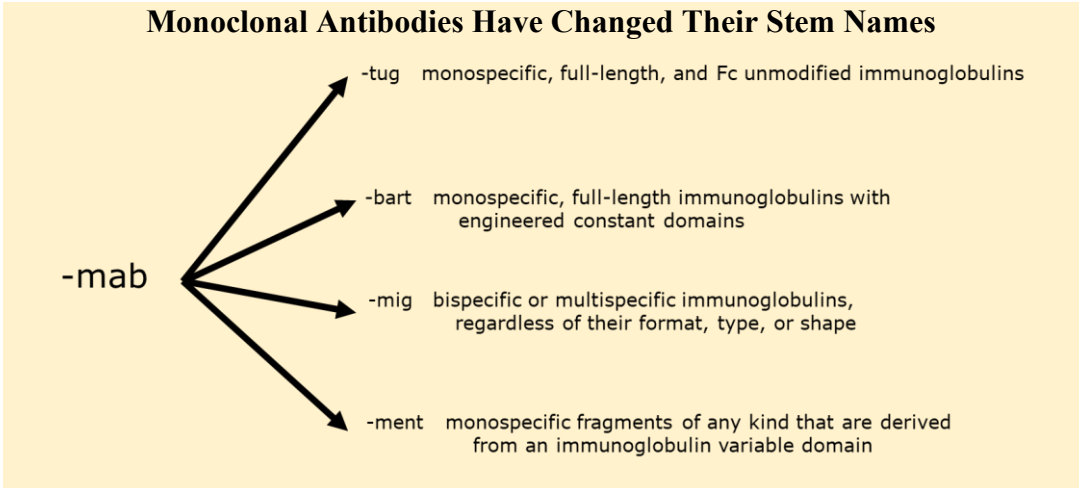
Updated USAN Stems for Biologic Drugs		
Stem	**Name of Group**	**Examples**
-rsen	Antisense oligonucleotides	mipomersen
-cel	Cell therapies	tisagenlecleucel
-ase	Enzymes	alglucerase
-poetin	Erythropoietin-type blood factors	epoetin alpha
-som	Growth factors	somatrem
-stim	Colony stimulating factors	filgrastim

Some companies attempt to be clever with the non-proprietary names of their drugs. For example, Bristol Myers Squibb's (BMS) drug asunaprevir, which also has its internal name BMS-650032, gets part of its name -sun- from the chemist who synthesized it: Li-Qiang Sun [105]. Dasatinib, another drug by BMS, has the internal name BMS-354825 and was named after an employee: Jagabandhu Das. Although not its inventor, he was the medicinal chemist at BMS whom the team felt was important enough to recognize for his influence [106].

Once a regulatory agency approves a drug for marketing, it is given a brand name. Brand names are proprietary names that other companies cannot use and are the company's exclusive property. Once a drug comes off patent, the brand-name still belongs to the company. The company proposes brand names to the regulatory agency, and they must be distinct from the non-proprietary name. Unlike USAN, the FDA will either accept or reject a brand-name and not propose alternatives. For example, acetaminophen is the non-proprietary name, but Tylenol is the brand-name. Phenytoin is the non-

Table 8 Examples of USAN Stems		
Stem	**Explanation**	**Example**
-clone	Hypnotic/tranquilizer	eszopliclone
es-	Dextrorotary enantiomer	eszopiclone
-zumab	Humanized monoclonal antibody	alemtuzumab
-ximab	Chimeric monoclonal antibody	infliximab
-previr	Serine protease inhibitor	telaprevir
-tu-	Targets tumor	trastuzumab, rituxumab

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



proprietary name, but Dilantin is the brand-name. Once the brand name is approved, the company can market the drug under that brand name until it is no longer sold in that country. Once off-patent, the drug may be marketed by either its brand or non-proprietary name. Furthermore, once the drug is off-patent, another company can develop its own brand-name version. For example, Janssen's infliximab is sold under the brand name Remicade. Once off-patent, Hospira, later acquired by Pfizer, started marketing a biosimilar to infliximab under the brand name Inflectra.

Brand names have always tried to be catchy and reflect some aspect of the drug. Although brand names may not have meaning, that hasn't stopped companies from being clever. Chlorpromazine, a drug used to treat schizophrenia and often used to sedate agitated patients quickly, was marketed as Thorazine because it was said to hit you like Thor's hammer. Furosemide was marketed as Lasix because it was said to last 6 hours. Sildenafil was marketed as Viagra because of its similarity to the word 'vigor.' Today, brand names are a big business. Whole companies are created to help drug companies name their drugs. Millions of dollars are spent on the process. Focus groups test names before a company approves them. However, getting the brand-name right and approved by the regulatory agency at the time of approval poses substantial financial risks, as marketing campaigns are typically planned for the day the drug is approved. If the regulatory agency rejects a brand-name, it could delay marketing by months, resulting in a corresponding loss of revenue.

Over time, the WHO updates its guidelines to reflect changes in drug classes and the development of new drug types. For example, in 2017, the WHO updated its guidelines for naming biologics, whose stems now include aspects related to their structure or function. In 2021, they recommended that the -mab suffix be used for monoclonal antibodies, which has been a staple of naming monoclonal antibodies for decades, with hundreds of drugs being named in this manner, be removed and replaced with four new stems: -tug, -bart, -mig, and -ment for antibody fragments [107]. Examples of drugs under the new naming system are tafolecitug, romosozubart, teclistamig, and abrezosment.

When the first biosimilar, Zarxio (filgrastim), was approved by the FDA in 2015, a 4-letter suffix was added to the end of the name, -sndz (filgrastim-sndz), to distinguish the generic name from the brand-name drug and its manufacturer. However, shortly after, the requirements changed again. The 4-letter stem must meet several criteria: it cannot be promotional, it cannot include abbreviations that might be misinterpreted as another drug, it must be unique to that product, and it must exclude numbers and lack meaning. Furthermore, the 4-letter suffix must be lowercase and consist of three distinct, non-proprietary letters. These new requirements prevent -sndz from ever being used again. In 2019, the guidelines were updated again, stating that the 4-letter suffix would be applied to both biosimilars and brand-name original products [108].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Patents and Exclusivity

Without patents, there would be no pharmaceutical industry. Patents are legally binding documents issued by a country that prohibit others from making, selling, marketing, or using the product for a specified period of time. Contrary to popular opinion, patents do not give *you* the right to make your product – they prohibit others from doing so. In the US, patents are issued by the US Patent and Trademark Office and have a 20-year term. While 20 years may seem like a long time, keep in mind that this period encompasses the entire drug development process, from the time the patent is issued to the time the drug is approved for marketing, which typically takes 10 years or more. It is not unusual for manufacturers to have less than 10 years actually to recoup their investment.

Patents, in effect, protect a company's intellectual property. Without a patent, any other company can market and sell the product developed by the originator company. There are many different types of patents, but patents for pharmaceuticals generally fall under utility patents, which are also called "patents for inventions"; these can be obtained on processes, machines, manufactures, compositions of matter, or a new and useful improvement, providing the invention meets three criteria:

1. Useful, also referred to as utility. The invention is useful if it *"provides some identifiable benefit and is capable of use."* It cannot be theoretical or frivolous and must have some credible or specific function.
2. Novelty. The product must be new and judged not to be obvious based on 'prior art', which is previous publications, patents, posters, or otherwise available to the public in some form. Essentially, the invention must be novel. There is a one-year grace period for the inventor to file a patent after releasing the invention to the public. For example, if a person presents a new drug at a scientific conference for the first time, they will then have 1 year to file for their patent. However, companies rarely take this risk and will not disclose their products until a patent has been filed.
3. Nonobviousness. It is not that hard to make a novelty claim; it is much harder to show nonobviousness. An invention is obvious if someone with ordinary skill in a specific field (in legal terms, they call this 'art') would find your invention obvious, given the current state of the field. If the invention is beyond the abilities of someone with ordinary skill in the art, then this meets the definition of nonobviousness. The inventor must demonstrate that the invention differs from previous ones.

The most important part of a patent is its claims, which are the legally binding statements regarding the scope or boundaries of the patent. Pharmaceutical companies can lay a host of patents around a new drug to protect its intellectual property. They can file patents for the chemical structure of the drug (these are composition of matter patents), how the drug is manufactured, chemically synthesized, or created (these are process patents), the formulation of the manufactured product, including route of administration or if some drug delivery system is employed, or how the drug may be used to treat patients (these are Method of Use patents).

While the general life of a patent is for 20 years, this term can be extended through patent extensions. The FDA also grants sponsor exclusivity, a period during which generic competition is prevented from entering the market. Some of these are intended to entice manufacturers to study their drugs in populations that might not otherwise be studied, like new antibiotics can receive 5 years of exclusivity, drugs approved for use in children can receive pediatric exclusivity and have 6 months of exclusivity, and drugs approved for rare, orphan diseases can receive 7 years of exclusivity. Furthermore, already approved drugs may receive 3 years of exclusivity under certain circumstances, including when they have undergone human studies, such as a change in formulation from a capsule to a tablet, or when they are approved for an additional disease state. Some of these patent changes result from changes to patent law, such as the Drug Price and Competition and Patent Term Restoration Act of 1984, widely known as the Hatch-Waxman Act. Under this act, an inventor can file for a patent extension based on the duration of clinical testing, specifically the period between filing an Investigational New Drug (IND) application and drug approval, not to exceed 5 years, and with a

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

remaining patent period of no more than 14 years. Typically, sponsors receive the full 5 years.

After a patent has expired, generic competition may enter the field – theoretically. And this is where we delve into the dark side of patents. Pharmaceutical companies typically file a host of patents around a product. These are referred to as 'patent thickets' because they frequently overlap. They are designed to prevent competitors from entering the market and to extend a manufacturer's monopoly over a drug. NBC News reported that, on average, a pharmaceutical company will file more than 140 patents for a drug [109], almost three-quarters of which are filed after the drug is approved [110]. This approach, while not illegal, is certainly successful. One of the more notable extensions is Humira, a biological drug approved for the treatment of rheumatoid arthritis in 2002 and the world's number one selling drug for many consecutive years. AbbVie filed more than 311 patents, of which 165 were approved, to prevent biosimilar competition from entering the market. Furthermore, AbbVie has been aggressive in suing competitors trying to enter the market. Often, these lawsuits are settled out of court, in which the competitor agrees not to market their generic for a specified period. These tactics held off the entry of Humira biosimilars for more than 20 years, until 2023.

Another approach to stave off generic entry is through reformulation [111]. Tricor (fenofibrate) was approved in 1998 for the treatment of high cholesterol. Once the exclusivity period expires, a generic manufacturer can file an Abbreviated New Drug Application (ANDA). The inventor can file a lawsuit challenging the ANDA as patent infringement, which, by law, gives the court 30 months to resolve the case. This allows the inventor to have up to 30 more months of marketing time. If, after 30 months, the FDA approves the ANDA "at-risk," knowing that the legal system may rule against the generic drug and have to pay damages to the inventor. In 2000, a generic competitor filed an ANDA against Tricor, and Abbott filed suit. During litigation, Abbot developed another formulation of Tricor, let's call it Tricor-2, which changed the drug's concentration. Abbott then removed Tricor-1 from the market. Within 6 months, Tricor-2 had 97% of the market. After the generic Tricor-1 was approved and the lawsuit settled, it could not compete with Tricor-2 because it could not be substituted for It due to differences in drug strength, and Tricor-1 was subsequently discontinued for sale. Hence, Abbott (the manufacturer) defended their drug against competition for years. In 2004, they repeated the process, creating Tricor-3, which differed from Tricor-1 and Tricor-2 in that Tricor-3 could be taken with or without food, whereas Tricor-1 and Tricor-2 required it to be taken with food. These reformulations successfully staved off competition for more than 7 years.

Patent life can also be extended through use patents, which extend the overall patent life by applying the invention to different diseases. An example of this is alemtuzumab, which was first granted a patent in

Patent Exclusivity vs. Regulatory Exclusivity

Patent exclusivity and regulatory exclusivity are means to protect products from competition, but do so through different mechanisms.

- **Patent exclusivity** is typically for 20 years and is granted by the US Patent and Trademark Office (PTO). It prevents others from making, using, or selling your drug. It is granted only for drugs that are novel, non-obvious, and useful. Secondary patents may be used to extend the 20-year exclusivity period.
- **Regulatory agencies grant regulatory exclusivity** and can vary by type. For example, orphan drug designation can provide 7 years of exclusivity, whereas pediatric exclusivity is 6 months. Regulatory exclusivity exists independently of patent exclusivity and can provide protection when patent exclusivity does not. Whereas patent exclusivity protects the invention itself, regulatory exclusivity protects the data and market approval rights generated through clinical testing.

A drug may have both patent and regulatory exclusivity, which can overlap. The one that expires later determines when a generic can enter the market.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

1998 for the treatment of leukemia (U.S. Patent No. 5,846,534). Campath (alemtuzumab) was approved in 2001 for the treatment of B-cell chronic lymphocytic leukemia. In 2000, a patent was granted for the use of alemtuzumab in the treatment of multiple sclerosis (MS, US-6120766). This was followed by another use patent in MS in 2008 (US 2008-0267954). Campath was removed from the market in 2012 and, in 2019, all patents related to the antibody itself expired. However, the use of patents in MS remains, and was approved in 2014 under the brand name Lemtrada, which is expected to remain in place until around 2028. With the removal of Campath from the market, this step limited the development of biosimilars because a biosimilar has to be switchable from a brand-name product. If there is no brand-name product, there can be no biosimilar.

While it might seem as if the generic drug companies are getting taken advantage of, they are far from angels themselves. After the Hatch-Waxman Act took effect in 1984, some generic companies nearly destroyed the industry before it got started. Generic drugs also receive exclusivity. To promote the entry of generics into the market, the first generic drug applicant to file an ANDA challenging a brand-name drug is eligible for 6 months of exclusivity from the FDA. The first generic to enter the field can receive a lot of money during this exclusivity period. Bolar Pharmaceuticals quickly filed 40 ANDAs after the act took effect, aiming to be the first to enter the market for several brand-name drugs. The problem was that all of these ANDAs contained fraudulent data. The FDA's review policy at the time was "first in, first out." It was found that this policy was not being followed. Further investigation revealed that some FDA reviewers were being bribed to give preferential review to specific companies, and that some of the ANDAs from these preferred companies also contained fake data stolen from the brand-name drug NDA. In total, 39 individuals and nine companies pleaded guilty to the charges or admitted their culpability. These episodes were collectively known as the 'generic drug scandal' and plunged confidence in generics for decades [112].

Where does this put consumers? Patents are necessary for pharmaceutical companies to maintain the incentive to innovate and create new products. Still, at the same time, there are many examples of companies gaming the system to squeeze out as much profit as possible for as long as they can, leaving consumers waiting for cheaper generic drugs. In contrast, insurance companies and governments are left holding the bag. The ability of pharmaceutical companies to manipulate the system is one of the reasons why pharmaceutical companies are seen and presented as evil by the lay press and media.

Non-US Regulatory Authorities

This chapter has focused on the FDA and the regulations in the United States. Every country has its own regulatory authority responsible for approving drugs. Europe has the European Medicines Agency (EMA); Japan has the Pharmaceuticals and Medical Devices Agency (PMDA); China has the National Medical Products Administration (NMPA); and the United Kingdom has the Medicines and Healthcare Products Regulatory Agency (MHRA). Each of these countries has its own rules and regulations governing the approval and sale of drugs within its borders. For example, in the US, the FDA evaluates and approves a drug, whereas in the EU, the Committee for Medicinal Products for Human Use (CHMP) issues an opinion. The European Commission then formally issues the Marketing Authorization. Although countries have harmonized on many activities, such as submissions using the Common Technical Document, approval in one country does not guarantee approval in another. For example, two liposomal formulations of doxorubicin, a cancer drug, were approved in Europe but not in the US. There may be situations in which a drug is approved for one indication in one country but not for the same indication in another. For example, Boniva (ibandronate) was approved in Europe for the prevention of skeletal events in patients with breast cancer with bone metastases, but it was only approved for osteoporosis in the US [113].

A movement is underway in which multiple regulatory agencies across countries review the same application simultaneously and exchange non-binding opinions to expedite the review process and bring drugs to patients more quickly. This effort, called Project Orbis, was launched in 2019 by the FDA's Oncology Center of Excellence. The first approval under this framework was for lorlatinib (Lorbrena), a treatment for ALK-positive lung cancer. The Australian Therapeutic Goods

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Administration and Health Canada participated in this review, enabling simultaneous approval in both countries. It should be noted, however, that each country issued its own label; a harmonized label across all three countries was not issued. As of 2023, more than 80 oncology drugs have been approved using Project Orbis. Although this program is still in its infancy, it is expected that more approvals will be issued under this framework in the future and may be applied to other divisions, not just oncology.

Summary

The development and approval of a drug is a complicated dance between the drug company and regulatory bodies. It's kind of funny the amount of mind-reading that happens at a company regarding what the FDA will think about a particular issue. The Regulatory Affairs department in a drug company provides strategic advice on regulatory issues, interacts with regulatory authorities, and prepares compliant, approvable submissions worldwide.

The regulations for drug development approval are constantly changing. There is a desire to expedite and improve the drug approval process, getting drugs to patients faster; everyone recognizes how slow, difficult, and expensive it is. And yet, regulatory reform in one country might not occur in another, and drug development then proceeds according to the slowest, most onerous link in the chain, if a company wants global approval. Greater harmonization is needed worldwide if we truly want to bring drugs to patients faster. Still, unfortunately, the current political climate in many countries is in no mood for greater globalization of drug development and approval.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Chapter 4

Principles of Pharmacokinetics

At a minimum, three conditions must be met for a drug to have an effect:

1. The drug must get to the site of action;
2. The drug must engage its target, either by activating or blocking its natural effect; and
3. The target must do "something."

The three conditions are referred to as *The Three Pillars* (Figure 17). If a drug is taken orally but its site of action is in the brain, and it doesn't reach the brain, it won't work. If the drug reaches the site of action but fails to engage its target, it won't be effective. Lastly, if the drug binds to its target but does so weakly or at a low level of response, its effect may not be noticeable.

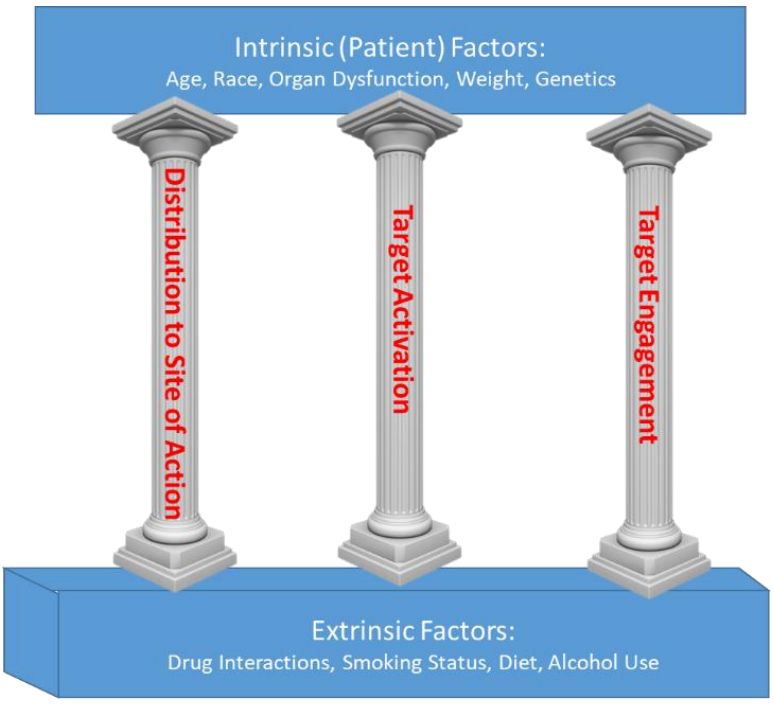


Figure 17. The Three Pillars are necessary for a drug's effect, as well as the intrinsic (patient-specific) and extrinsic factors that affect them.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Physicochemical Properties of Drugs

To understand how drugs are absorbed by diffusion across a cell membrane, such as in the gastrointestinal tract, and how this applies to any cell membrane in the body, it is essential to understand a drug's physical and chemical properties. These properties are also important for the drug to interact with its receptor and have a pharmacologic effect. These properties include solubility, logP, pKa, and molecular weight. These will be discussed in turn.

- Solubility refers to how easily the drug is dissolved in a solvent, whether that is blood, plasma, or gastrointestinal fluids. Solubility is often dependent on the solvent's pH. Highly soluble drugs are absorbed more easily than poorly soluble drugs because drugs that are not dissolved in solution cannot cross membranes.
- LogP is a measure of a drug's lipophilicity, where P is its octanol-to-water partition coefficient. LogP is the logarithm of P. Octanol is like cooking oil. Adding octanol to water forms a partition, with octanol sitting on top of the water, much like when you mix oil and vinegar, such as in Italian dressing. If you now add a drug to this mixture, the drug will partition between the octanol layer and the water layer. The higher the LogP, the more lipophilic ("fat-liking") the drug is and the more nonpolar the compound is. A LogP of 3 means that for every molecule of drug in the water phase, 1000 molecules will be found in the octanol phase (10^3). For a drug to be absorbed orally, a rule of thumb states that its LogP should be < 5 [114]. However, a small logP (< 1) can indicate that the drug may have difficulty crossing a cell membrane, which is composed of fatty chains. A logP around 2 is ideal. LogP can also be dependent on the pH of the water layer. When the water layer has a pH of 7.4, logP is referred to as logD.
- pKa is a measure of how strong an acid a drug is. For example, a carboxylic acid can exist in solution as

  $$R\text{-}COOH \leftrightarrow RCOO^{-} + H^{+}$$

  The pKa is the pH at which 50% of the molecules in solution are dissociated (and in the R-COO^{-} state), and the other 50% are in the protonated state (and are in the R-COOH state). The smaller the pKa value, the more acidic the drug is. For example, a strong acid like hydrochloric acid has a pKa of < 0, whereas water is 15.7. Weak acids have a pKa between 2.5 and 7.5, whereas weak bases have values between 5 and 11. Acid drugs are generally protonated at pH values below their pKa and dissociated at pH values above their pKa. Predominantly uncharged drugs will tend to have higher absorption than charged drugs.
- Molecular weight is the sum of all the atomic weights of the drug, i.e., all the carbons, hydrogens, oxygens, etc. The unit of molecular weight is the Dalton (Da). For example, acetaminophen has the formula C8H9NO2, where the subscript indicates the number of atoms. Its molecular weight is 151 Da. The lower a drug's molecular weight, the easier it is for the drug to be absorbed. One rule of thumb is that the molecular weight of an orally absorbed drug should be < 500.

Pharmacokinetics[5] deals with Pillar 1 and, simply put, is the study of "what the body does to the drug." Pharmacokinetics studies the absorption, distribution, metabolism, and elimination (ADME) of a drug in the body. Pharmacokinetics is not the same as pharmacy; pharmacy refers to the preparation and dispensing of drugs, a field distinct from pharmacokinetics.

[5] This chapter and the next are the most technical in the book. It's impossible to talk about pharmacokinetics without a few equations. I have tried to keep these to as few as possible. This chapter is also the most science-y of the book, so please slog through it.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

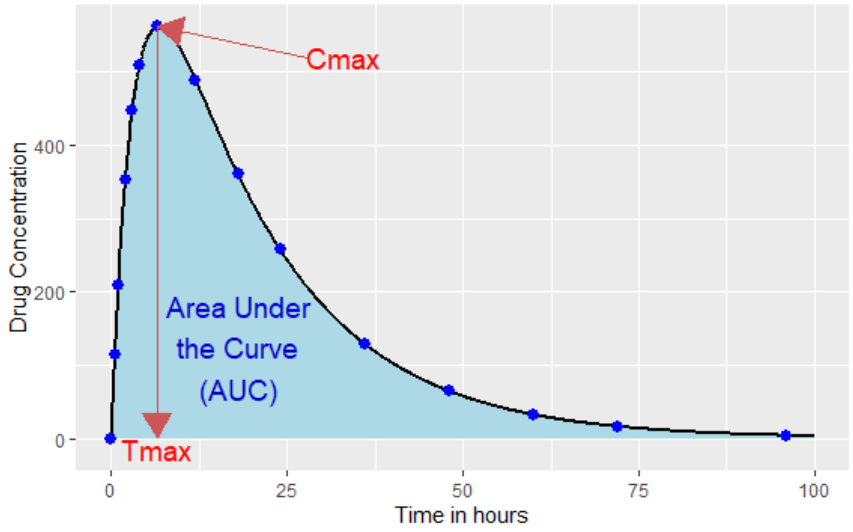


Figure 18: A representative concentration-time profile after oral administration for a drug having a half-life of 12 hours. Also shown are the three exposure measures: AUC, C_{max}, and T_{max}.

This chapter will discuss the basic principles of pharmacokinetics and some factors affecting a drug's pharmacokinetics. The next chapter on Principles of Pharmacodynamics will deal with the last two pillars, target engagement and activation, and will discuss the basic principles of pharmacodynamics, or "what the drug does to the body". Before we begin, a note about this chapter. Most of the drugs we take are small molecules that are taken orally – think aspirin and Tylenol. As such, this chapter will focus on the pharmacokinetics of orally administered small molecules. When something deviates from this theme, it will be highlighted.

Pharmacokinetic Measures of Exposure

Before discussing pharmacokinetics, a brief digression into pharmacokinetic measures of exposure is needed. When a variable is measured, like the distribution of weight among males in a population, the results are often summarized using descriptive statistics. These statistics, such as the mean, median, minimum, and maximum, aim to describe the data using a few measures that capture its essence. The same is true for pharmacokinetic data, where the concentration-time profile is summarized into a few measures of "exposure," or the "how much" drug the body is exposed to. Three of these are particularly important:

- The maximal concentration (C_{max}) is the highest observed concentration after dosing and can be considered similar to the maximum in descriptive statistics.
- The time that C_{max} occurs is called T_{max}.

Important Pharmacokinetic Parameters of Drugs

- Exposure – how much drug the body is "exposed" to. The two primary parameters used to characterize exposure are the area under the curve (AUC) and the maximum concentration (C_{max}).
- Clearance – the volume of drug removed from the blood per unit time; has units of volume/time.
- Volume of distribution – the hypothetical volume a drug must distribute into to obtain a given drug concentration.
- Half-life – the time it takes for 50% of a drug to be removed from the body.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

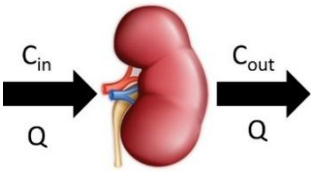
The smaller the T_{max}, the more rapidly the drug is absorbed. The longer the T_{max}, the slower the drug is absorbed.

- Area under the curve (AUC) is an integrated measure of overall exposure over time. It can be thought of as a weighted sum of all observations.

These concepts are illustrated in Figure 18. Cmax and T_{max} measure the rate and extent of absorption, while AUC measures the extent of absorption. Researchers can use these summary measures for many purposes. One way is to compare the exposure of a drug alone and in combination with another drug. If the AUC and C_{max} of both drugs remain unchanged when combined, it is safe to conclude that there is no interaction between them. But if the AUC and C_{max} are higher for one or both of the drugs when the two drugs are given in combination, then one may conclude that there is a drug interaction between the drugs. Or, if Cmax and Tmax change when food is administered, we may conclude that food affects the drug's absorption.

The Concept of Clearance

Clearance is a central concept in pharmacokinetics, possibly the most important, and can be defined in several ways [115]. In physiology, clearance refers to the volume of blood completely removed of a substance per unit time and has units of volume per time. Clearance, as defined in this manner, has its roots in kidney physiology, where physiologists were interested in removing waste products, such as creatinine, from the kidneys. Blood flows into the kidney at rate Q while the concentration of the substance in the blood, which could be a drug or some natural waste product in the body, enters the kidney with some value C_{in}. Blood leaves the kidney at the same flow rate, Q, but some of the drug is removed from the blood (into the urine) as it passes through the kidney, leaving the kidney with a concentration of C_{out}. This is shown below:



The organ clearance of the substance can be mathematically expressed as:

$$CL = \left(\text{Blood flow}\right) \times \left(\begin{array}{c}\text{fraction of drug} \\ \text{extracted by organ}\end{array}\right) = Q \times E = Q\left(\frac{C_{in} - C_{out}}{C_{in}}\right). \qquad (1)$$

In this case, Q is the blood flow to the kidneys, but if we were referring to the liver, Q would be the blood flow to the liver. The *extraction ratio* of a compound (E) can be defined by:

$$E = \left(\begin{array}{c}\text{fraction of drug removed} \\ \text{from the organ as it passes through}\end{array}\right) = \frac{C_{in} - C_{out}}{C_{in}} = 1 - \frac{C_{out}}{C_{in}}. \qquad (2)$$

But what exactly does Eq. (1) mean? Suppose a substance is efficiently cleared by the organ (such substances are called high extraction substances), then C_{out} would be zero, and E would be 1. Hence, CL ≅ Q, which means that clearance equals blood flow to the organ. Thus, the clearance of the substance is *flow-limited*; as soon as the blood reaches the kidney, the kidney completely clears the substance from the blood. This sets the upper limit on what CL can be – it can never exceed Q, the organ's blood flow. Now, suppose that the substance is not cleared by the kidney at all (such drugs are called low-extraction substances). In this case, E = 0 and CL = 0. In this case, the drug's clearance depends solely on the extraction ratio, regardless of blood flow. Such substances are called *extraction-limited*. Hence, organ clearance can range from 0 to Q. This range helps put an observed organ clearance value into perspective.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Organ clearance also has the useful property of being additive. In other words, if both the liver and kidneys can remove a substance from the blood, then the total systemic clearance from the body is the sum of kidney and liver clearances:

$$CL = CL_{kidney} + CL_{liver} . \quad (3)$$

In this manner, clearance represents the body's ability to remove a substance.

Clearance can also be defined in terms of pharmacokinetic principles. In a linear system whose properties do not change over time (more will be discussed on what this means later in the chapter), and assuming a drug that is completely absorbed by the body, the AUC of the drug is proportional to the administered dose:

$$AUC \sim Dose . \quad (4)$$

In other words, the drug's exposure is proportional to the dose, i.e., the higher the dose, the higher the AUC. This proportionality can be removed by introducing a constant, which is 1/clearance. Hence, Eq. (4) can be expressed as:

$$AUC = \frac{1}{Clearance} Dose \quad (5)$$

which can be rearranged to give one of the fundamental properties of pharmacokinetics:

$$Clearance = \frac{Dose}{AUC} . \quad (6)$$

This equation assumes complete drug absorption into the body. For most drugs, this is an unrealistic assumption. This assumption can now be relaxed by introducing the concept of bioavailability (which will be discussed in more detail later). Suffice it to say that bioavailability, denoted as F, is the proportion of the dose that reaches the systemic circulation and ranges from 0 to 1, where 0 is when none of the drug reaches the systemic circulation, and 1 is when all of the drug reaches the systemic circulation. Eq. (6) can then be relaxed and made more general by:

$$Clearance = \frac{F \times Dose}{AUC} . \quad (7)$$

Since the amount of a drug absorbed is often not known, Eq. (7) is rearranged to:

$$\frac{Clearance}{F} = \frac{Dose}{AUC} \quad (8)$$

where CL/F in Eq. (8) is referred to as *apparent clearance* and CL as defined in Eq. (6) as *total systemic clearance*. If the dose is given orally, CL/F is defined as the *apparent oral clearance*. Hence, with apparent clearance, clearance is confounded with the proportion of drug absorbed, but with total clearance, it is assumed that all the drug is completely absorbed.

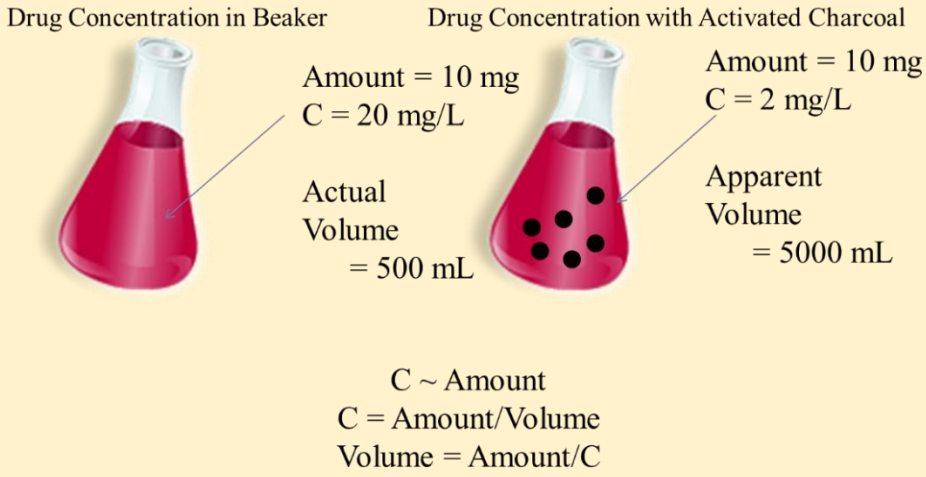
Eq. (8) can also be rearranged in terms of AUC:

$$AUC = \frac{F \times Dose}{Clearance} . \quad (9)$$

Hence, if the AUC changes for any reason while the dose remains constant, then the change in exposure can be attributed to either a change in bioavailability or a change in clearance. This property is useful in explaining changes in exposure seen with two drugs that have a metabolic drug-drug interaction.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Volume of Distribution Illustrated



$$C \sim Amount$$
$$C = Amount/Volume$$
$$Volume = Amount/C$$

The volume of distribution can be understood using a drug dissolved in a beaker of water. Suppose you have a beaker with 500 mL of water and add 10 mg of the drug. Now, you measure the drug concentration in the water and find it is 20 mg/L. Since concentration is mass/volume, and you added 10 mg to the beaker, the volume of distribution must be 0.5 L (10 mg/20 mg/L).

Then, suppose you repeat the experiment, but this time, you add activated charcoal to the beaker. Activated charcoal binds to the drug, reducing its concentration in solution. This time, the drug concentration in the beaker is 2 mg/L. The apparent volume of distribution is 5 L. We say the volume is 'apparent' because the only way to get a concentration of 2 mg/L with 10 mg of added drug is if the volume is 5 L. In humans, the same concept applies.

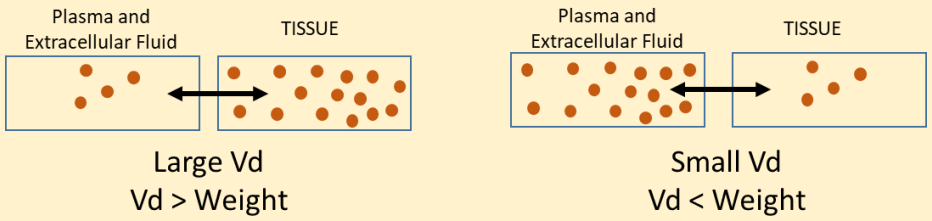
The Concept of Volume of Distribution

The second important concept in pharmacokinetics is the volume of distribution (Vd). Suppose a drug like propofol, an anesthetic used in surgery, is given intravenously as a rapid bolus injection of 40 mg. Immediately after administration, the propofol concentration in blood is measured at 1 mg/L. Concentration equals mass over volume. In this case, the mass administered was 40 mg and the concentration was 1 mg/L. Hence, its volume of distribution must be 40 L. Now, suppose a drug is given intravenously as a 200 mg dose, and its measured concentration is 0.5 mg/L. Its volume of distribution is then 400 L. How is this possible? The human body is about 70 L in adult males (assuming a density of 1 kg/L). How is the drug's volume of distribution twice that of the average male?

Like the beaker example with activated charcoal, the body has its own binding elements. Plasma proteins, such as albumin and α1-acid glycoprotein, can bind drugs, effectively preventing them from being found free-floating and unbound in plasma. Furthermore, within tissues, intracellular proteins can bind drugs. Hence, these binding proteins increase the volume of distribution, sometimes much larger than the volume of the human body. For this reason, the volume of distribution in humans is called the *apparent volume of distribution*. It is said to be apparent because it bears little or no resemblance to a physiological or actual volume. It's a mathematical construct, not physiologic or anatomic in nature, that relates a dose to a concentration.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drugs with Large V_d Have High Tissue Penetration



Drugs with a large volume of distribution (Vd) have high tissue penetration and bind to intracellular macromolecules. Drugs with low V_d mostly stay confined to the plasma and extracellular fluid and have low tissue penetration.

The apparent volume of distribution can also indicate how the drug is distributed throughout the body. Drugs with a volume of distribution in the 3 to 10 L range, like monoclonal antibodies or warfarin, are confined to plasma and intracellular fluid, whereas drugs with large volumes of distribution, like isavuconazole with its 400 L volume of distribution or chloroquine with a volume of distribution of 15,000 L, are highly tissue-bound. The volume of distribution may not tell you where a drug is distributed to, whether it's the heart or muscle, but it can tell you where it's not. For instance, monoclonal antibodies have a small volume of distribution, indicating that their distribution is largely confined to the extracellular space outside cells. Hence, monoclonal antibodies are generally ineffective against targets inside cells; their targets are either on the cell membrane or in the bloodstream.

The Concept of Half-life

Half-life is perhaps the easiest concept in pharmacokinetics because its name is self-explanatory. Half-life is the time it takes for 50% of a drug to be removed from the body (Figure 19). If a drug has a half-life of 1 day, then half the dose remains one day after drug administration, once all the drug has been absorbed into the body. What's interesting about half-life is that it's cumulative: 50% remains after one half-life, 25% remains (50% of 50%) after two half-lives, 12.5% remains (50% of 25%) after three half-lives, etc. Generally speaking, after 5-6 half-lives, the drug is essentially removed from the body (as a problem, can you compute how much remains in the body after six half-lives?).

Mathematically, half-life (T1/2) has an interesting relationship to clearance and volume of distribution:

$$T1/2 = Ln(2)\frac{V_d}{CL}. \quad (10)$$

Hence, half-life is directly proportional to both clearance and volume of distribution. If V_d increases, half-life increases; if CL increases, half-life decreases. Now we can estimate what might happen if two drugs interact and one of them, called the perpetrator drug, slows the clearance of the other drug, called the victim drug, by half. The victim drug's half-life will double. But can the same be said in reverse? If a victim drug's half-life doubles in the presence of the perpetrator drug, can it be said that it's because the perpetrator drug slows the clearance of the victim drug? No, because the change could be due to a change in clearance, a change in volume of distribution, or a change in both. Interestingly, there are cases where perpetrator drugs affect both clearance and volume of distribution to an opposite but equal extent, leading to no change in the half-life of the victim drug.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

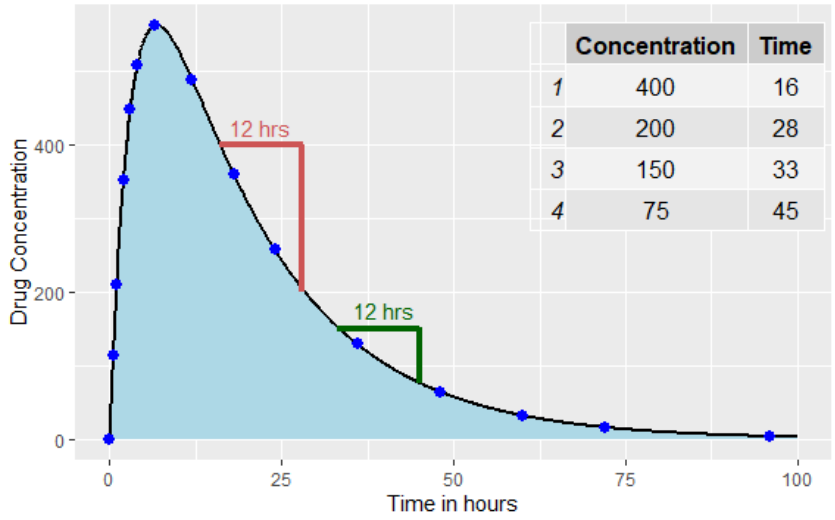


Figure 19: The Concept of Half-Life. This is a representative concentration-time profile after oral administration, showing the concept of half-life. The box indicates the drug concentration at the time indicated. The drug's half-life is 12 h; drug concentrations decline by half every 12 hours during the drug's elimination phase (after T_{max}), regardless of where you are on the curve after the peak has occurred.

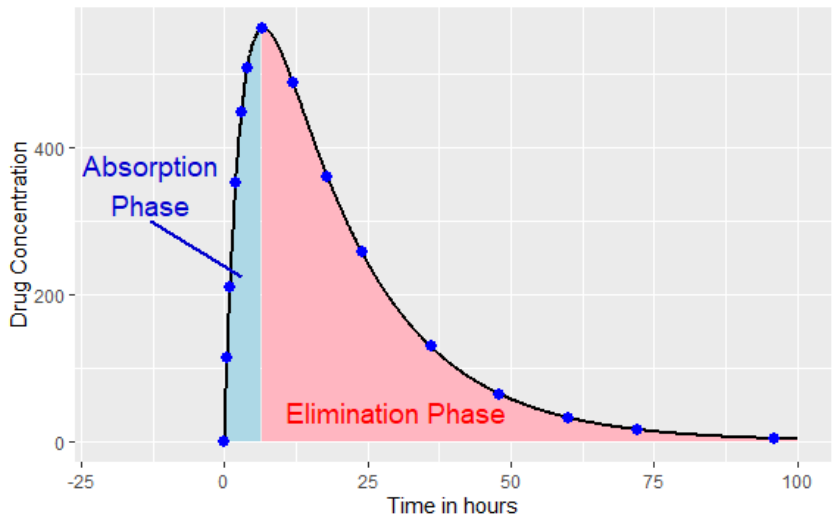


Figure 20: After extravascular administration, where dissolution from the formulation is rapid, drug concentrations increase and then decline. Concentrations increase when the rate of absorption exceeds the rate of elimination and decline when the rate of elimination exceeds the rate of absorption. At its peak, known as T_{max}, the rate of absorption equals the rate of elimination.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drug Absorption

When drug therapy is initiated from an extravascular site of administration (e.g., oral, rectal, intramuscular, buccal, or sublingual), and drug concentrations are measured in blood, concentration-time profiles (like those above) increase to a peak and then decline. Unless another dose is taken, drug concentrations will decline to zero over time due to the drug's half-life. But why do drug concentrations increase and then decrease? It concerns the competing forces of absorption and metabolism/excretion. For drugs in which dissolution from the formulation is rapid and immediate, drug concentrations increase when absorption of the drug into the body predominates; the rate of drug absorption into the blood from the extravascular site of administration is greater than the rate of metabolism and excretion from the blood (Figure 20). When the rate of drug absorption equals the rate of drug elimination, drug concentrations will be at their peak. After that, the rate of drug metabolism and excretion from the blood exceeds the rate of drug absorption from the extravascular site of administration, and concentrations decrease over time. For drugs administered intravenously, there is no absorption phase because the drug is delivered directly into the bloodstream, and concentrations will usually always decrease over time.

For some drugs, an interesting and uncommon phenomenon called flip-flop pharmacokinetics can occur. In this situation, the rate of release of the drug from the formulation is slow relative to absorption, i.e., elimination is faster than absorption [116]. Although it has been noted for some immediate-release drugs, particularly those that have low intestinal membrane permeability making them difficult to absorb, flip-flop is more often reported with extended or sustained-release formulations. An example is the extended-release (ER) formulation of Campral (acomprosate) (Figure 21), a drug used to maintain alcohol abstinence. Volunteers were administered the extended-release formulation three times daily for 7 days and then, 17 days later, received a 15-minute intravenous infusion of acamprosate [117]. The ER formulation showed different pharmacokinetics compared to intravenous administration; the half-life was particularly longer with oral administration, which was highly suggestive of flip-flop. Some non-oral routes of administration, such as intramuscular or subcutaneous administration, may also show flip-flop. If flip-flop is not recognized, a drug's pharmacokinetics may be misinterpreted. In the case of acamprosate, if intravenous data were unavailable, the half-life may have been misreported as 33 hours instead of 3 hours.

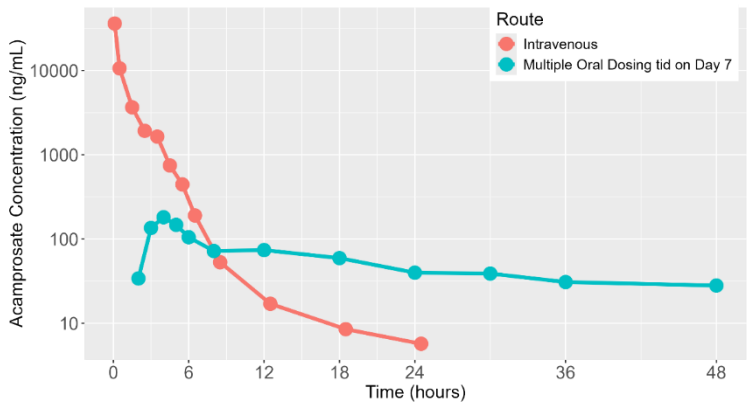
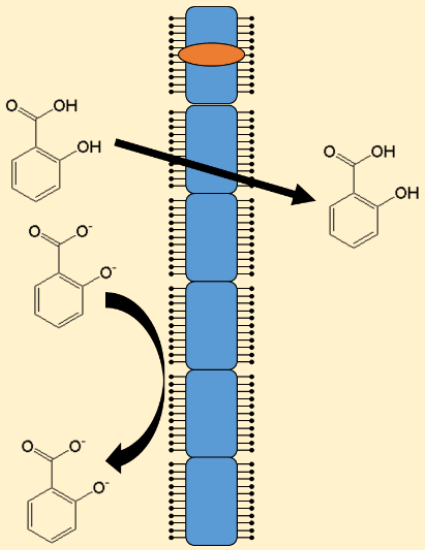


Figure 21: Example of flip-flip pharmacokinetics, where the rate of elimination is faster than the rate of absorption throughout the entire time interval. A total of 24 healthy volunteers were enrolled in a crossover study and administered Campral (acamprosate). Volunteers received 666 mg (333 mg × 2) of enteric-coated acamprosate tablets three times daily (tid) for 7 days, and following a 17-day washout, received a single 15-minute intravenous infusion of 666 mg acamprosate. The half-life of the intravenous route of administration was 3 hours, whereas for the oral route it was 33 hours, a more than 10-fold increase. This increased half-life is a hallmark of flip-flop kinetics. Figure redrawn from [117].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Two physicochemical properties of a drug that influence its absorption potential are solubility (how readily the drug will dissolve in water) and permeability (how easily the drug crosses cell membranes, whether by diffusion or transporter-mediated) [118]. Drugs with high solubility and permeability tend to be rapidly absorbed and are not affected by food. Such drugs have very short T_{max} (on the order of 1 hour). Drugs that have high solubility and low permeability usually take longer to absorb and have reduced exposure when the drug is coadministered with food. Conversely, drugs with low solubility and high permeability also usually take longer to absorb but have higher exposure when the drug is coadministered with food. Drugs with low solubility and low permeability have erratic absorption, making it difficult to predict how food will affect absorption. Solubility and permeability also indirectly affect peak drug concentrations (C_{max}). Permeability indirectly affects the rate of drug absorption. The faster a drug is absorbed, combined with how much of the drug is absorbed and the drug's volume of distribution, all contribute to C_{max} and T_{max}.



Only unionized (uncharged) drugs can pass through biological cell membranes

Drug absorption from an extravascular site to the blood is complex. Since oral administration is by far the most common route, this will be our focus (Figure 22), although the general processes hold for other extravascular routes. First, a tablet or capsule has to disintegrate. This may occur in the stomach for immediate-release formulations, but disintegration occurs in the intestine for some sustained-release formulations. Next, the disintegrated particles must dissolve in the fluids of the stomach and the gastrointestinal (GI) tract. This is why taking a pill with a full glass of water is a good idea. Whether a drug dissolves in a solution depends

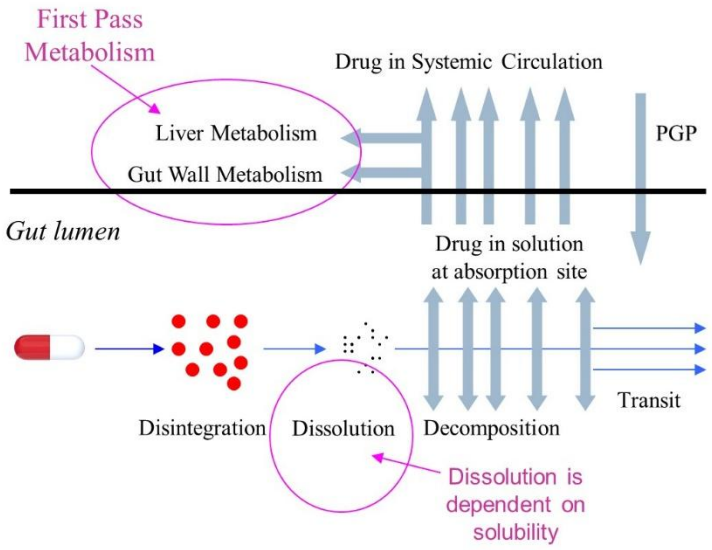


Figure 22: Schematic of the Processes Involved in Oral Drug Administration. Legend: P-gp, p-glycoprotein.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

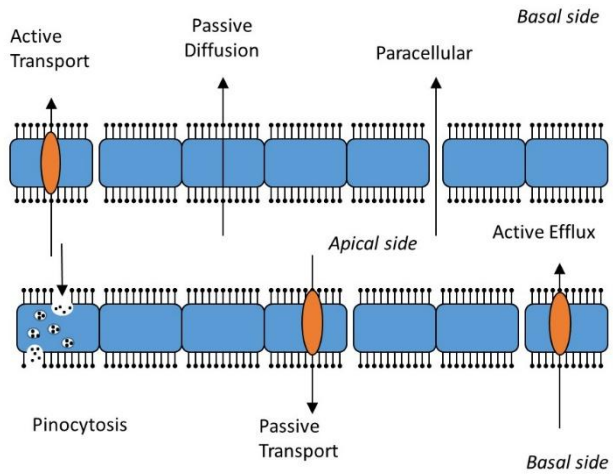


Figure 23: Schematic Showing the Routes by Which a Drug Can Be Absorbed from the Small Intestine into the Blood. Active transport requires a protein (shown in orange in the figure) to help cross the membrane. Diffusion across the cell membrane occurs directly and depends on the drug's size and charge.

on its physicochemical properties and the volume of the fluid. The larger the fluid volume, and depending on the excipients, the more likely the drug will go into solution, and more of the drug itself will go into solution. The drug will not be absorbed if the formulation does not disintegrate or dissolve.

For most drugs, absorption occurs in the small intestine, not the stomach, as many commonly believe. The stomach is primarily a holding and churning tank whose primary function is to break down food into a form that can pass into the intestine. Once in the small intestine, a couple of things can happen to the drug:

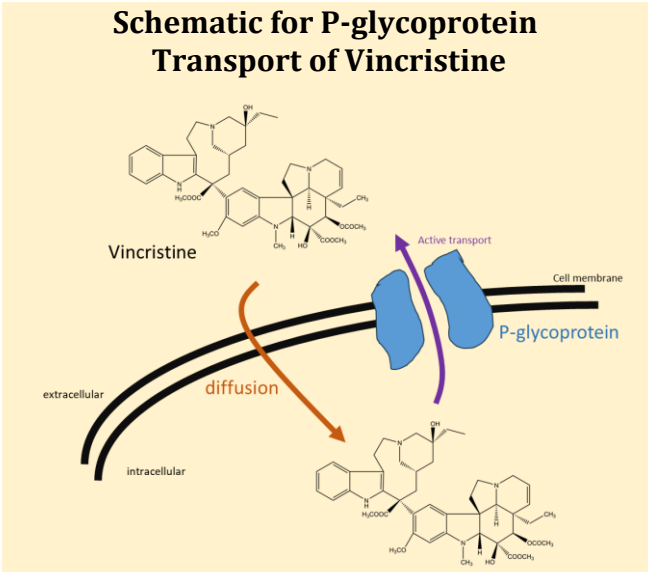
- It could transit through the small intestine and not get absorbed, ultimately passing into the feces.
- Within the small intestine are human and bacterial enzymes that can metabolize the drug. The metabolized drug could then be absorbed or transit through the intestine into the feces.
- The drug can attempt to pass through the gastrointestinal (GI) cells on its way to the blood and be absorbed.

Passing through the cells of the GI tract from the inside (apical side) to the blood (basal side) is not easy, and there are many possible routes. Suppose the drug is sufficiently nonpolar (think oily). In that case, it can cross the lipid membrane on the apical side of the cell, passively diffuse to the other side, and then cross the membrane on the basal side of the cell, where it is now in the blood, leaving the GI tract. Only unionized drugs can cross cell membranes by diffusion. Drugs like salicylic acid (the original version of aspirin) had low absorption because the pH in the small intestine was such that the acid was too ionized to cross the cell membrane. Acetylsalicylic acid (the modern form of aspirin) was developed as a non-ionized form of salicylic acid with better absorptive properties.

Another way a drug could enter the bloodstream is through pinocytosis, where the cell engulfs the drug by forming an invagination and then absorbs it via small vesicles, which are subsequently released on the other side of the cell. Literally, the cell "eats" the drug and absorbs it. The drug could also squeeze between cells and enter the blood by a paracellular mechanism. So maybe if drugs were designed to be very nonpolar, absorption would be easy. If only this were the case. Along the cell surface lining the GI tract is a water barrier that prevents very nonpolar drugs from crossing the lipid membrane. Hence, a drug must be sufficiently nonpolar to cross a lipid membrane but not too nonpolar so that it can't pass through the unstirred water layer. These concepts are illustrated in Figure 23.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Within the last ten years, the role of transporter proteins in drug absorption has become apparent. Transporters are proteins that span the cell membrane, bind the drug, and transport it across the membrane from one side to the other. Transporters can work in both directions, as either influx or efflux transporters, and can be active (require energy) or passive (do not require energy) in nature.



P-glycoprotein (P-gp) is a major transporter that limits the absorption of many cancer drugs. Once the drug is absorbed passively into the cell, P-gp pumps it back into the GI lumen, thereby limiting its absorption.

A major GI tract transporter is the efflux p-glycoprotein (P-gp), which is extensively distributed throughout the body's organs, including the small intestine, kidney, liver, and brain. P-gp pumps back out absorbed drugs from inside to outside the cell. P-gp acts on countless drugs, as well as natural products like lipids, steroids, and peptides. Another important transporter, an efflux transporter like P-gp, is breast cancer resistance protein (BCRP), which has been shown to prevent the absorption of many cancer drugs and is co-localized with P-gp in the same tissues, but is particularly expressed in the placenta to protect the fetus from harm. Not all transporters of clinical interest are efflux transporters; the organic anion transporter (OAT) transports many drugs from outside to inside the cell, including statins used to treat hypercholesterolemia.

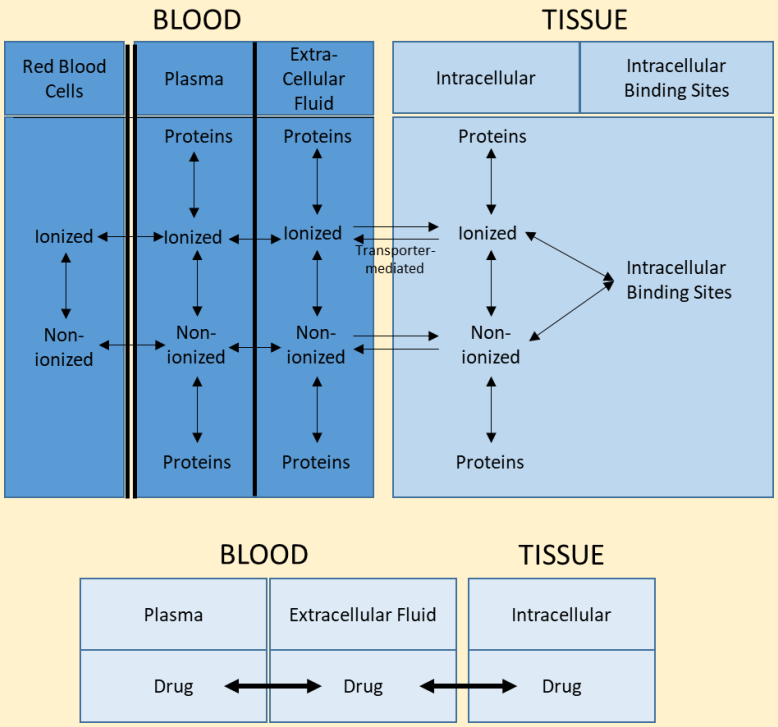
Further, to prevent drug absorption, bacterial and human enzymes in the gut attempt to metabolize the drug as it passes through. The most important of these is cytochrome P450 3A4 (CYP3A4). More about the CYPs will be discussed later in the chapter. For now, let's say that CYP enzymes metabolize nonpolar drugs, making them more polar, thereby potentially inactivating the drug and rendering it more difficult for the metabolite to be absorbed. This metabolism is referred to as intestinal first-pass metabolism.

Based on the distribution of P-gp and CYP3A4 enzymes in the GI tract and the fact that the genes encoding these proteins are located on the same chromosome, it is believed that these two proteins have evolved in a complementary manner to protect the body against foreign substances. CYP3A4 expression is highest in the duodenum, steadily declining when moving towards the colon. On the other hand, the expression of P-gp is the opposite – it is highest towards the rear of the small intestine and lowest near the duodenum. The net effect is that, as the drug moves into the small intestine, it encounters high levels of P-gp, which pumps it back from inside GI cells into the lumen. Then, as the drug moves through the GI tract towards the colon, it will encounter increasing levels of CYP3A4, which will attempt to metabolize it. This dynamic interplay results in decreasing the absorption of xenobiotics.

But once a drug is absorbed, it doesn't mean it has overcome all its challenges in reaching the site of action. Indeed, all the blood from the GI tract flows into the hepatic portal vein, which in turn supplies the liver, where the drug can be metabolized by liver enzymes —a process known as hepatic *first-pass metabolism*. The sequential process of the intestinal and hepatic first-pass is called first-pass metabolism. When a drug is given intravenously, it bypasses first-pass metabolism and becomes completely available to the body.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Schematic Showing the Reversible Distribution of Small-Molecule Drugs from Blood Compartments to Tissue Compartments



Top: Non-ionized drugs cross cell membranes using diffusion, transporter, or paracellular mechanisms. Bottom: Schematic showing the reversible distribution of biologics between blood and tissue. Due to their large size and ionic charge, many biologics have limited tissue distribution, with most remaining primarily in the plasma and extracellular fluid. Interstitial fluid, the fluid that fills the space between blood and tissues, is not shown.

Bioavailability

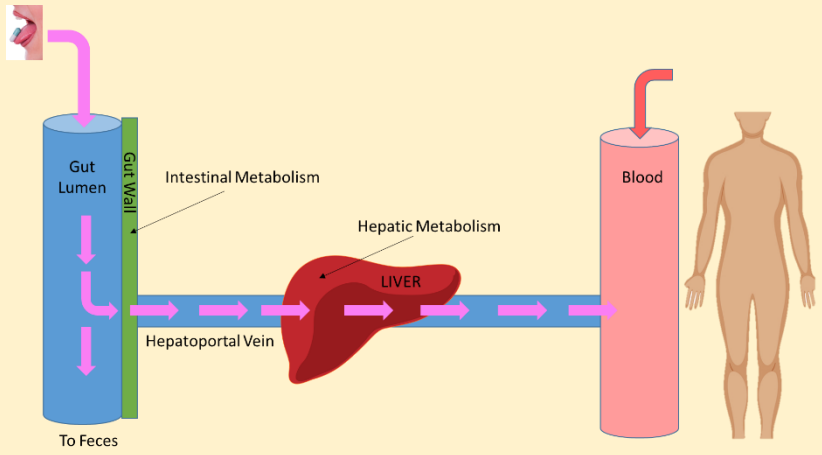
A central concept in absorption from extravascular sites is bioavailability, which can be absolute or relative. *Absolute bioavailability* refers to the proportion of drug absorbed after extravascular administration relative to intravenous administration. Drugs administered intravenously are not subject to first-pass metabolism and are completely bioavailable to the body's organs; therefore, the entire administered dose is available to the body. In contrast, extravascular drugs are subject to first-pass and may have reduced bioavailability compared to intravenous drugs. For a drug given at the same dose both intravenously and extravascularly, the absolute bioavailability (denoted F) of the drug is given by:

$$F = \frac{AUC_{ex}}{AUC_{iv}} \tag{11}$$

where the right-hand side of the equation is the ratio of the AUCs of the two routes of administration. F ranges from 0 to 1, where 0 means the drug is not absorbed at all and 1 means the entire dose is absorbed. Absolute bioavailability is the product of all the fractions of the drug escaping metabolism

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Illustration of First-Pass Metabolism



First-pass metabolism refers to the sum of all metabolic processes that occur before a drug is absorbed into the bloodstream after extravascular administration and distributed throughout the body.

and elimination during the first-pass. For instance, suppose 80% of a drug escapes gut/intestinal metabolism, is absorbed into the portal vein, and then 60% escapes liver metabolism. For some drugs, the lungs and enzymes in the venous blood can metabolize drugs, but in this example, assume 95% of the remaining drug reaches the site of action. Hence, the absolute bioavailability of the drug is 0.80 × 0.95 = 0.76, or 76%.

The other type of bioavailability, *relative bioavailability,* is used when comparing two extravascular routes of exposure. In this case, relative bioavailability can be calculated as the ratio of the AUCs between the two routes. Relative bioavailability is important with different formulations. For instance, a tablet may have 85% of the bioavailability of a capsule due to incomplete disintegration and the drug's inability to dissolve in gastrointestinal fluids.

Food can affect both the rate and extent of absorption of some drugs and improve the tolerability of drugs that may be difficult to take on an empty stomach [118] (Table 9). Food can have no effect, a positive effect, or a negative effect on drug absorption (Figure 24). Food tends not to affect the bioavailability of drugs with high solubility and high permeability and may also increase their absorption rate. These drugs include acetaminophen, caffeine, codeine, ketoprofen, and diltiazem.

Table 9

Effect of Food on Bioavailability Based on the Drug's Solubility and Permeability

	High Solubility	Low Solubility
High Permeability	F ↔ T_{max} ↑	F ↑ T_{max} ?
Low Permeability	F ↓ T_{max} ↓	F ? T_{max} ?

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

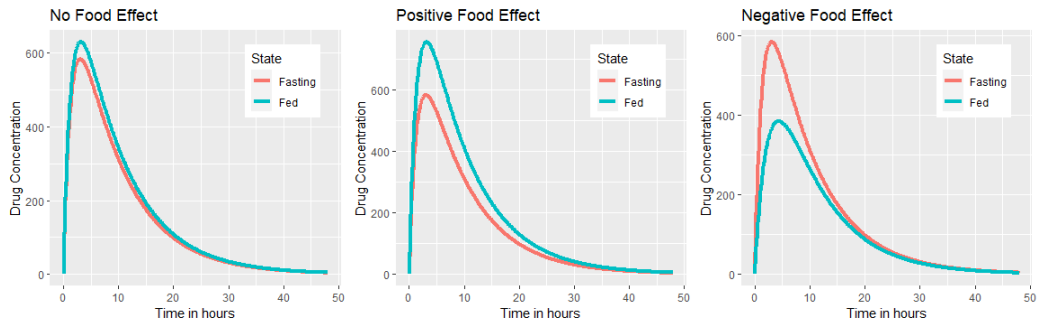


Figure 24: Possible effects of food on a drug's pharmacokinetics. Food may have no effect on drug concentrations (left), increase them (middle), or decrease them (right).

Food tends to increase the bioavailability of drugs with low solubility and high permeability and may have a variable effect on their absorption rate. These drugs include atorvastatin, ibuprofen, indomethacin, naproxen, and warfarin. Foods with high solubility and low permeability tend to reduce drug bioavailability and slow drug absorption. These drugs include amoxicillin, erythromycin, fexofenadine, metformin, and penicillin. Lastly, food has a variable effect on drugs with low solubility and permeability. These drugs include digoxin and ciprofloxacin. Does this mean that you should always take a drug with or without food? Not necessarily. As seen in the next chapter and the chapter on drug interactions, it depends on the drug's therapeutic index. Antacids and herbal supplements can also affect the rate and extent of absorption; however, this topic will be further discussed in the chapter on Drug Interactions.

A discussion of absorption would not be complete without considering the influence of formulation. Have you ever wondered why there are so many different types of aspirin available at the pharmacy? Orally administered aspirin is available in various forms, including immediate-, sustained-, delayed-, and modified-release tablets, chewable tablets, buffered tablets, effervescent tablets, and possibly others I am not aware of. But why are there so many? Some, such as buffered and effervescent, are designed to improve the safety profile and reduce gastrointestinal side effects, while others aim to alter how quickly and how much aspirin is absorbed. Changing the formulation can only affect how fast the drug is absorbed and how much of the drug is absorbed (Figure 25). The formulation will not change distribution parameters such as systemic clearance or volume of distribution.

Drug Distribution

The human body is a recirculatory system (Figure 26). Arterial blood flows to the organs, passes through them, and leaves as deoxygenated venous blood; it then flows to the lungs to be oxygenated, and the cycle begins again. Since the human body contains about 5 L of blood and cardiac output is about 5 L/min, it only takes a few minutes for a drug absorbed into the venous circulation to be distributed and equilibrated throughout the body. A drug's volume of distribution, denoted V_d, is a complex function of plasma volume plus the sum of all the other organ volumes corrected for the ratio of unbound drug in plasma to that inside the organ. For biologic drugs like monoclonal antibodies, the summation term is very small, and the volume of distribution is slightly larger than plasma volume. Hence, these drugs have volumes of distribution ranging from 5 to 8 L. For small molecules, Vd can be small or in the thousands of liters.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

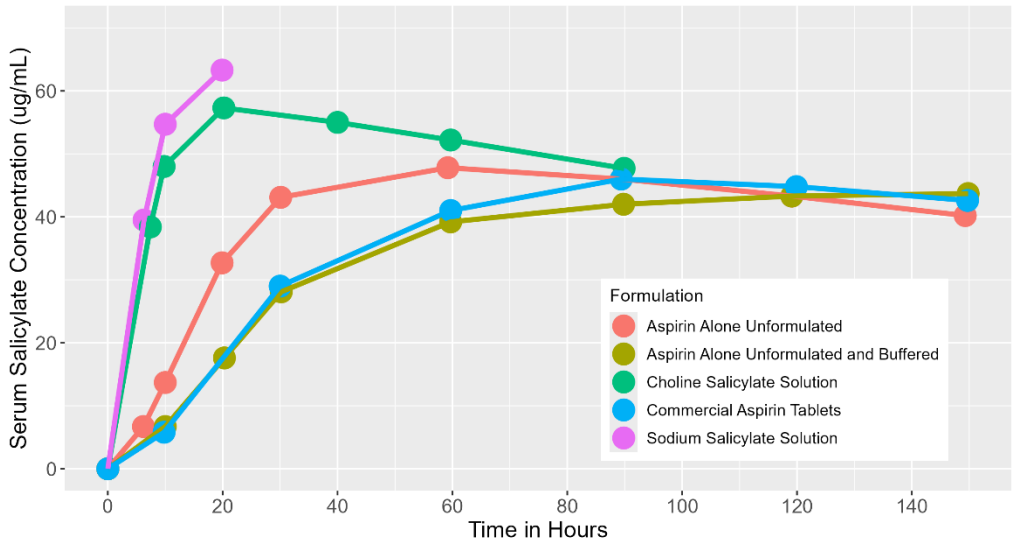
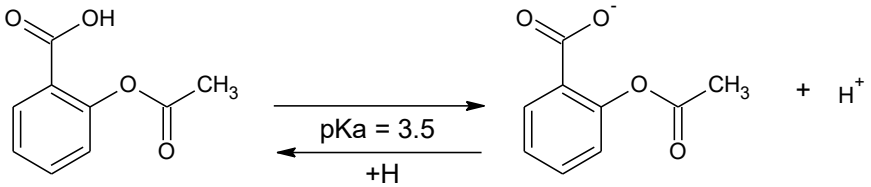


Figure 25: Salicylate concentrations after administration of different aspirin (sodium salicylate) formulations in healthy volunteers. Changes in formulation for non-intravenously administered drugs, such as oral administration, only affect the amount and rate of drug absorption. The faster the absorption rate, the higher the drug's peak concentration. Formulation does not affect disposition parameters, such as clearance or volume of distribution.

In any compartment, whether the stomach, small intestine, or blood, a dynamic equilibrium exists between the ionized and non-ionized forms of a drug, and how those forms distribute between blood components, such as plasma proteins and red blood cells, and tissue components, including proteins and DNA. For example, aspirin is a weak acid drug that can exist in two states, a nonionized and ionized form, as shown below:



nonionized form ionized form

A drug's pKa is the pH at which 50% of the drug exists in the ionized form. At the blood pH of 7.4, ~100% of aspirin is found in the ionized form. A drug placed at any point in the system will be distributed to the various sites until equilibrium is reached. The speed with which an organ equilibrates with plasma depends upon blood flow to the organ and the degree of permeability and vascularity of the organ. Higher blood flow and permeability contribute to higher equilibrium rates. For biological drugs, which are often proteins and significantly larger than small molecule drugs, distribution is more limited, typically confined to the blood, and has limited tissue penetration.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

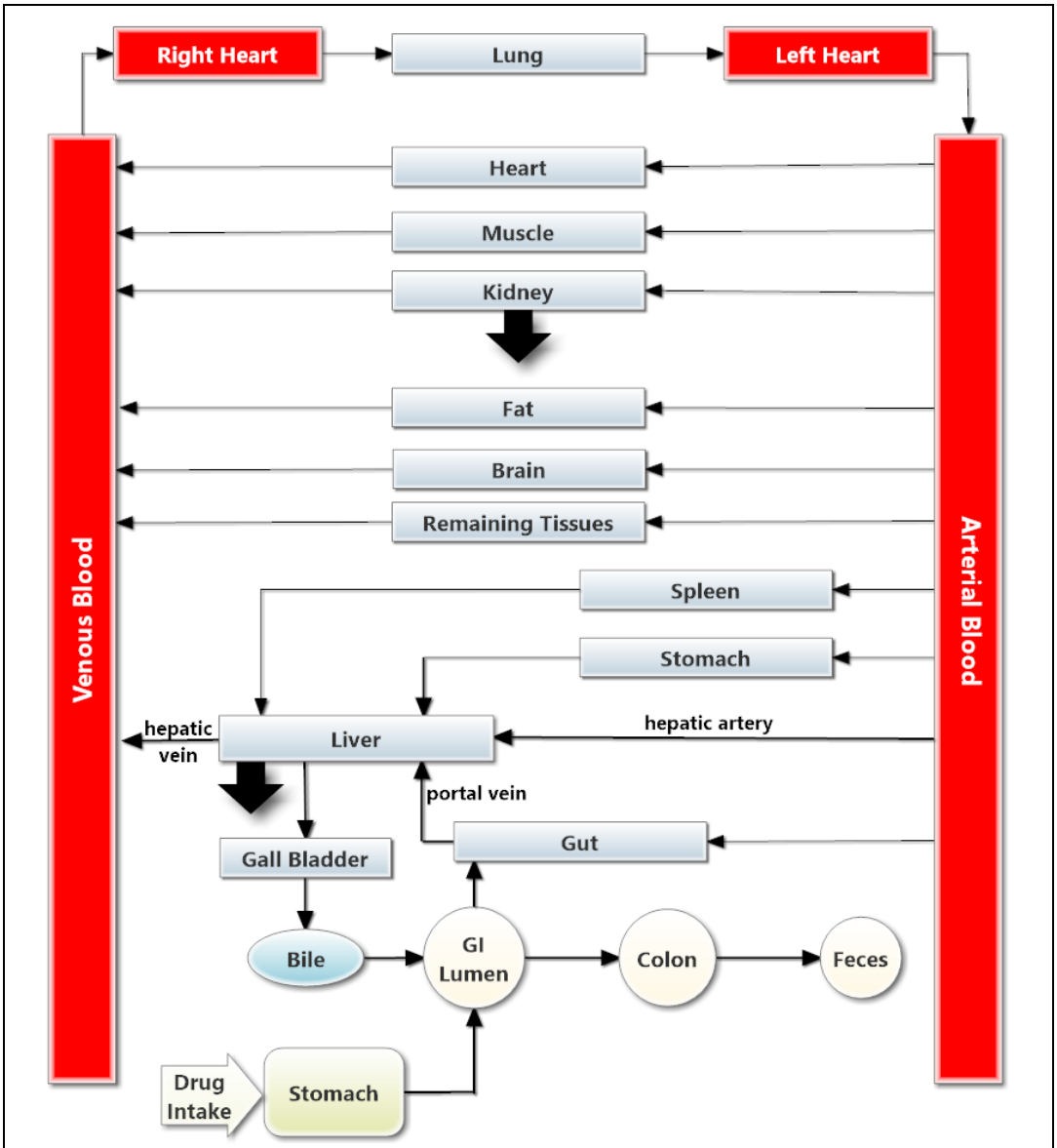


Figure 26: Schematic showing the recirculatory nature of the human body and major routes of elimination from the liver and kidneys. Fat arrows indicate routes of elimination. Also shown is how oral drug intake is processed.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

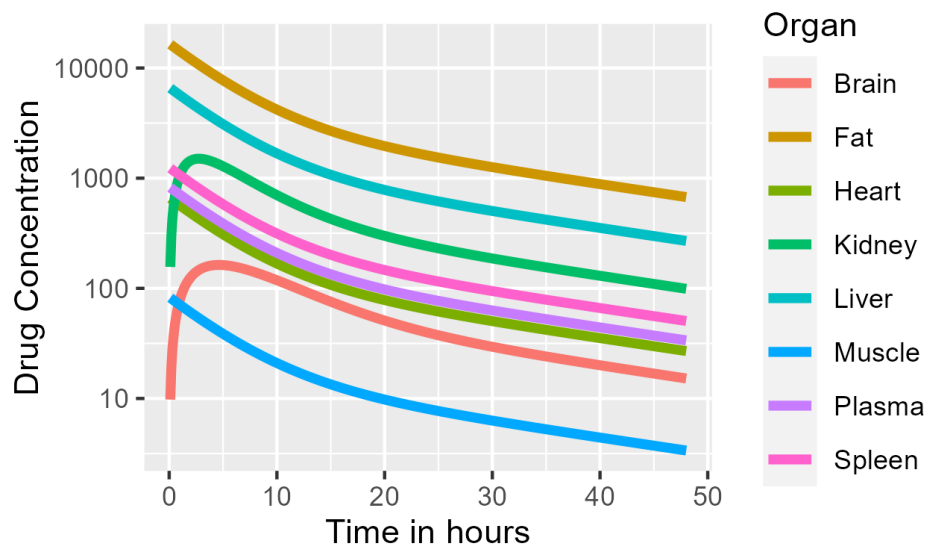


Figure 27: Drug concentration-time profiles in various organs after single-dose intravenous administration.

Small molecule drug concentrations in tissues reach equilibrium with plasma drug concentrations usually quite rapidly (Figure 27). Due to intracellular binding sites, organs, except for those with special barriers such as the brain, eye, and testes, tend to (but not always) have higher drug concentrations than plasma. The higher the drug concentration in the organ, the greater the drug binding to sites within the organ. Typically, the liver, spleen, and lung organs have the highest drug concentrations compared to plasma. If you compare the AUC of the tissue concentration-time profile to that of plasma, the ratio is called the tissue-to-plasma partition coefficient, denoted K_p,

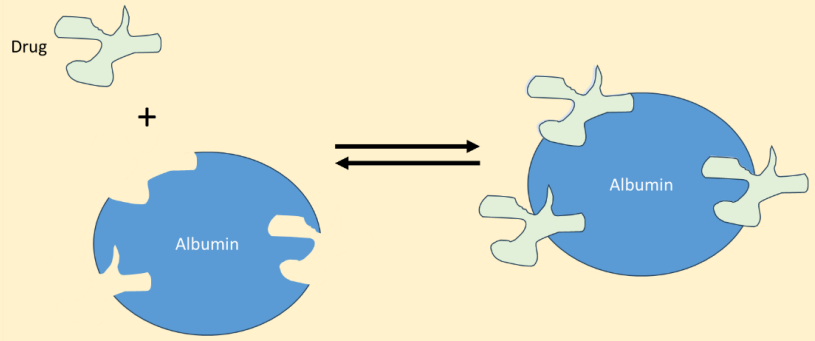
$$K_p = \frac{AUC_{organ}}{AUC_{plasma}} . \quad (12)$$

The larger the K_p, the greater the drug partitions into that organ. Due to their size, biologics are often confined to the vascular space, have a small volume of distribution, and exhibit K_p values < 1. The liver, spleen, and lungs tend to have high partition coefficients, whereas the brain tends to have a low one. It is also expected that poorly vascularized organs, such as fat, take longer to equilibrate than other organs, which is indeed the case.

Another important factor contributing to the distribution of small-molecule drugs is their binding to plasma proteins. The two major proteins involved are albumin and alpha-1-acid-glycoprotein (AAG). The total plasma protein content is approximately 7 g/dL, with albumin accounting for more than half, at an average concentration of 4 g/dL. Although its primary function is to help maintain plasma pH and osmotic pressure (which help maintain fluid in the capillaries), albumin also transports many hormones and fatty acids. Albumin also binds acidic and neutral drugs, mostly reversibly. Since only unbound drugs can distribute across membranes into tissues, bind to receptors, or be metabolized, drugs bound to albumin are shielded from these processes and act as a storage depot. Drugs that are very highly protein-bound (> 99%) have small volumes of distribution because they are retained in the plasma compartment.

The other important plasma protein that can bind drugs is AAG, an inflammatory protein that increases during inflammation. Although its concentration is lower than that of albumin (~0.1 mg/dL), it has a much higher affinity for binding drugs. Hence, albumin is said to have a high capacity

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Proteins like albumin and AAG can transport drugs and other proteins throughout the body. Albumin is not drawn to scale.

and low affinity for binding drugs, whereas AAG has a low capacity and high affinity. Because AAG has a low capacity, binding is saturable, meaning that if drug concentrations are sufficiently high, all binding sites will be occupied, and no further binding can occur. In contrast, albumin binding appears to be nearly limitless, as albumin concentrations are significantly higher than those of the drugs.

Disease states that affect protein binding, such as hypoalbuminemia, or inflammatory states, like those following myocardial infarction, increase AAG. Some diseases, like renal disease, can accumulate waste products (uremia) that displace drugs from their binding sites. And some drugs can displace other drugs from their binding sites. The effect of these changes and their clinical relevance have been debated in the literature. For years, it was thought that displacement of protein binding would lead to exaggerated drug effects, which would not be a problem in many cases but would be for narrow therapeutic drugs like warfarin. Today, this belief has been largely dispelled. Under most circumstances, changes in protein binding appear to have little effect. Only highly protein-bound drugs (> 99%), which have high intrinsic clearance (which will be discussed in the section on Drug Elimination), and are dose-titrated, appear to be affected by changes in protein binding [119]; this is a very small list of drugs.

Drug Metabolism

Major Sites of Drug Elimination From the Body are the Liver and Kidney, but can also occur in the Gastrointestinal Tract (both by GI cells and Microbiome), Lungs, Brain, Skin, and Blood.

Once absorbed, most drugs are substrates for various chemical reactions that change their chemical structure, physical properties, and biological effects. These new structures are called metabolites. Generally speaking, metabolites are more polar and less biologically active (though not always) than their parent structures. Sometimes, metabolites can be toxic and pose a danger to the body. The objective of metabolism is to increase the polarity and water solubility of the parent because, as discussed in the Drug Elimination section, these features will make it more likely that the metabolite is excreted from the body. Although the liver is the primary site of drug metabolism, the gastrointestinal tract also contributes significantly. Other sites of metabolism include the kidney, brain, and plasma.

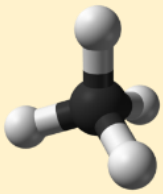
The body can perform many chemical reactions on a drug, which are too numerous to discuss here and are divided into two types: Phase I and Phase II reactions. Phase I reactions are typically oxidations, reductions, or hydrolysis reactions, usually performed by the cytochrome P450 (CYP) isozyme system. In contrast, Phase II reactions add large functional groups to the parent compound, a process known as conjugation. It was mentioned earlier in the book that logP is a measure of a drug's polarity. Drug metabolism often decreases a compound's LogP, which indicates the metabolite is more polar than the parent compound.

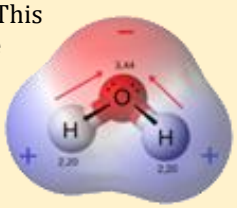
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Polar, Nonpolar, and Amphipathic

Much is discussed in this book about a drug's polarity and whether a drug is polar, nonpolar, or amphipathic, because polarity can affect every pharmacokinetic (PK) process, such as absorption and distribution, as well as pharmacodynamic (PD) processes, including drug binding to its receptor. Many drugs are molecules consisting of atoms (usually carbon, oxygen, nitrogen, and hydrogen) held together by covalent and ionic bonds. Covalent bonds are chemical bonds between two atoms that share electrons, while in an ionic bond, one atom donates its electrons to the other. Ionic bonds can be thought of as polar bonds in which one atom accepts electrons from the other.

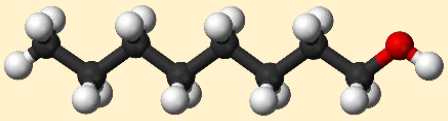
Covalent bonds can be polar or nonpolar. Nonpolar covalent bonds have an equal sharing of electrons between atoms; each atom contributes one electron to the shared pair equally. Methane (CH_4, see image) is completely nonpolar as each hydrogen shares electrons equally with the carbon. Polar covalent bonds do not share electrons equally. One atom exerts its pull on the other atom's electrons, leading to an unequal sharing of electrons. Water (H_2O) is polar (see image); it has an asymmetrical structure and does not share its electrons evenly between the oxygen and hydrogen. This results in the oxygen atom having a slightly negative charge and the hydrogens having a slightly positive charge, creating an electrical dipole similar to that found in a battery. Hence, one part of the molecule has a slightly positive charge, and the other part has a slightly negative charge.





Polar molecules consist of atoms that have predominantly polar and/or ionic regions, while nonpolar molecules consist of atoms that are mostly nonpolar. Drugs with lots of oxygen, fluorine, chloride, nitrogen, sulfur, or ionic bonds tend to be polar. Drugs with many carbons and hydrogens tend to be nonpolar; fats and oils are nonpolar. This is why oil and water do not mix; fats are nonpolar, and water is polar.

Amphipathic, also known as amphiphilic, molecules possess both polar and nonpolar regions, thereby exhibiting both polar and nonpolar properties. A useful one is that they can dissolve in both water and nonpolar solvents. As such, they are used as detergents, soaps, emulsifiers, and antibacterial agents. The image on the left shows 1-octanol, which consists of a long carbon chain and an oxygen atom (in red). Octanol is technically amphipathic but acts as a nonpolar substance.

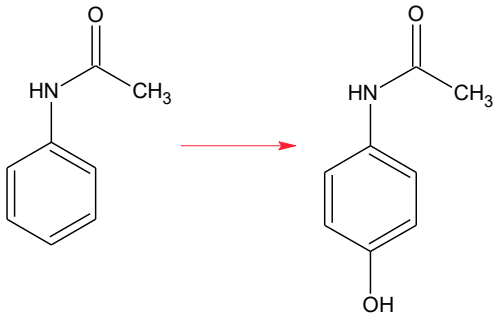


pH also plays a crucial role in determining the polarity of ionic compounds. Ionic compounds dissociate in solution and are charged. But under the right conditions, they may retain their counterion and act as nonpolar species. For example, carboxylic acids (-COOH) may dissociate to COO^- and H^+. However, when there is a preponderance of H+ ions, such as at low pH, the equilibrium may shift the acid back to its COOH form, making it nonpolar. In contrast, bases like amines (NH2-) in an acid may become ionized to NH3+, making them polar.

Figures reprinted from Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

An example of a Phase I oxidation reaction is the addition of -hydroxyl groups (OH) to a drug. This is exemplified by the metabolism of acetanilide, an early analgesic drug no longer in use:



Acetanilid p-hydroxyacetanilid

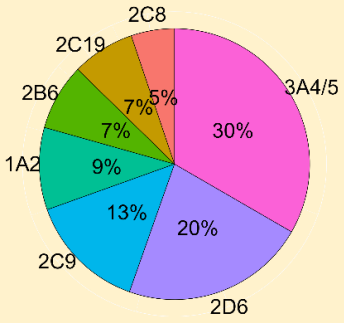
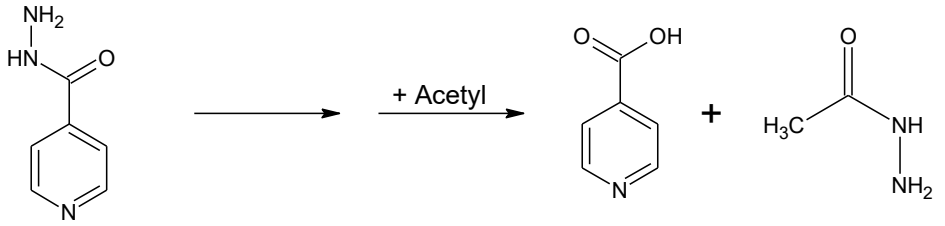


Figure modified from Figure 1 of [120].

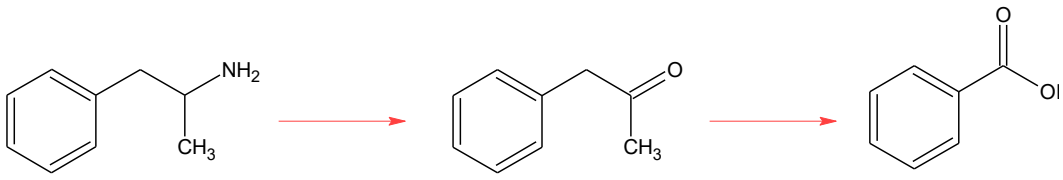
The LogP of acetanilide is 1.7, while that of the hydroxylated metabolite is 1.35, indicating that the metabolite is more polar.

An example of a Phase I hydrolysis reaction is the metabolism of isoniazid, an antibiotic used to treat tuberculosis. Isoniazid is first acetylated (a Phase II reaction that will be discussed shortly) and then hydrolyzed into isonicotinic acid and acetylhydrazine. The latter metabolite is an example of a toxic metabolite, one that is harmful to the liver and can be potentially life-threatening.



Isoniazid Isonicotinic Acid Acetylhydrazine

Lastly, an example of a Phase I reduction reaction is the deamination of amphetamine into phenylacetone. Amphetamine has a LogP of 1.8, while that of phenylacetone is 2.1, so this is one of those instances where the metabolite is less polar than the parent. However, phenylacetone is almost immediately oxidized into benzoic acid, which has a LogP of 1.8.

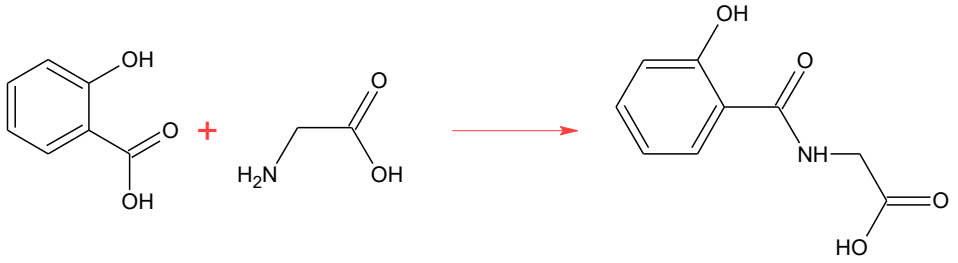


Amphetamine Phenylacetone Benzoic Acid

Phase II metabolism adds endogenous substances to the drug, such as glucuronic acid, sulfate, acetate, methyl groups, or amino acids, most often resulting in a loss of pharmacologic activity. These products are often called conjugates, have a higher molecular weight than the parent compound, are highly water-soluble, and can be excreted in the urine. For example, salicylic acid's

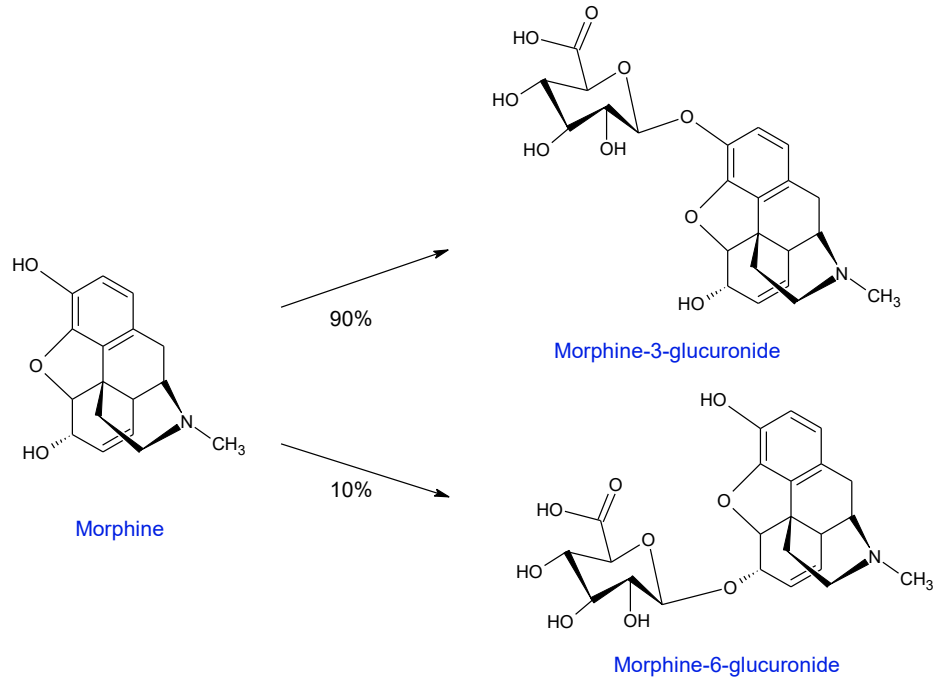
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

major metabolite is the glycine conjugate:



Salicylic Acid Glycine Salicylic Acid Glycine Conjugate

Although most conjugates are inactive, one exception is morphine, whose 3-O-glucuronide and 6-O-glucuronide conjugates are the dominant active forms of the drug:



Morphine illustrates an important concept: drug metabolism and its pathways can be quite complex. Often, drugs are metabolized into many Phase I and II metabolites. For example, Figure 28 shows the metabolism of mirabegron, a drug used to treat an overactive bladder. Although after oral administration, almost half the dose of mirabegron is excreted unchanged in the urine, 10 metabolites were identified in urine, of which eight were also found in plasma. These metabolites were formed through Phase I and II reactions, which involved hydrolysis, glucuronidation, oxidation, and deamination.

The major enzyme system responsible for drug metabolism is the Cytochrome P450 (CYP) system. CYPs are responsible for most Phase I drug reactions and are found in the endoplasmic reticulum membrane in the liver, small intestine, brain, kidney, lung, and skin. Physiologically, CYPs are responsible for the formation of numerous natural products, including steroids and hormones. CYPs are responsible for metabolizing hundreds of drugs. Cytochromes are proteins that contain iron; the P450 designation originates from the peak of the UV spectrum (450 nm), which was first used to identify them in the 1950s. Genes encoding the CYP enzymes are identified by an abbreviation followed by a number indicating the gene family and a letter for the subfamily, such as CYP3A. Sometimes, another number is added to indicate the individual gene, as in CYP3A4. CYP3A is the major CYP isozyme responsible for drug metabolism, followed by CYP2D6.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

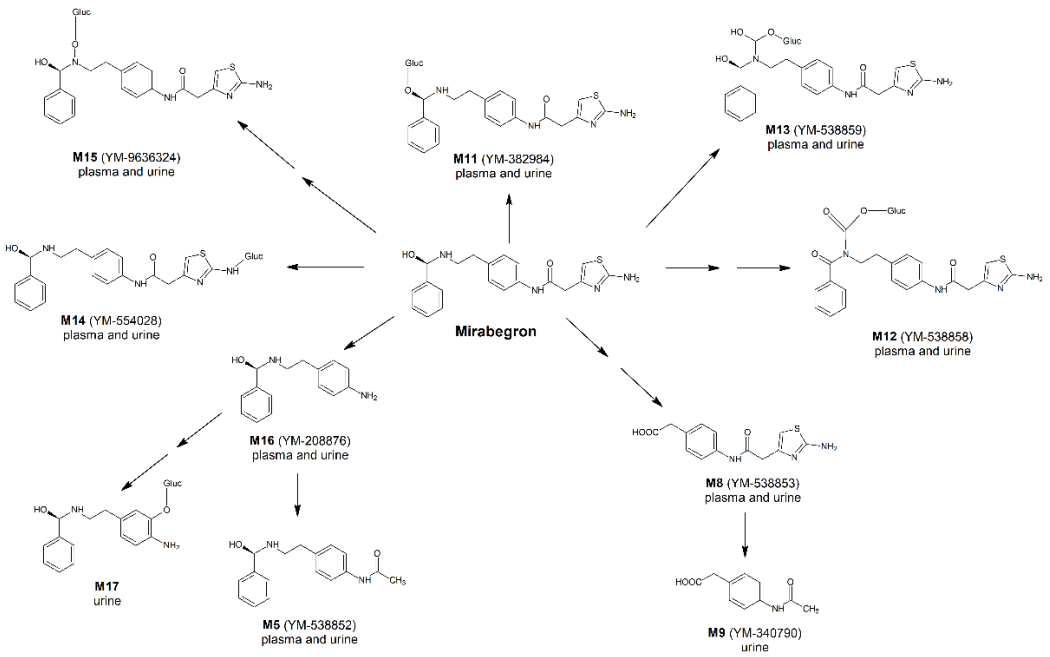


Figure 28: Proposed metabolic pathways of mirabegron in humans. Mirabegron (shown in the middle) is metabolized to several Phase I and Phase II metabolites. Gluc indicates glucuronide metabolite. Metabolites found only in urine are indicative of kidney metabolism. Figure courtesy of Astellas Pharma.

Some of the CYP isozymes have interesting properties. CYP3A4 and CYP1A2 are inducible, meaning that certain drugs, such as rifampin or anti-epilepsy medications like phenytoin or phenobarbital, increase the expression of these CYP isozymes, leading to a greater number of enzymes available for metabolism and faster drug clearance. Some drugs, like carbamazepine, induce their own metabolism, a process known as autoinduction. Some isozymes, like CYP2D6, CYP2C19, and CYP2C9, are polymorphic, meaning that at the genetic level, there are individual differences in DNA structure that lead to functional differences (called the phenotype) in the isozyme's activity. For example, 5-15% of Caucasians, 0-5% of Africans, and ~1% of Asians have reduced CYP2D6 activity and are considered "poor metabolizers" of CYP2D6 substrates.

Interestingly, ~2% of Caucasians, ~30% of Ethiopians, and 1-2% of Asians harbor multiple gene copies of CYP2D6, leading to increased enzyme expression and classification of these individuals as "ultra-rapid metabolizers" of CYP2D6 substrates. The remaining population consists of "extensive metabolizers," "intermediate metabolizers," and "poor metabolizers." These polymorphisms can result in different dose requirements for certain drugs, such as nortriptyline, where ultra-rapid metabolizers typically require 250 to 500 mg daily, extensive metabolizers require 100 to 250 mg, and poor metabolizers require only 20 to 50 mg. The discovery of these polymorphisms led to the establishment of a new field, pharmacogenetics, which examines their impact on drug pharmacokinetics and pharmacodynamics. Today, the field is known as pharmacogenomics, which accounts for the vast amount of genomic data available to an individual. Lastly, other drugs, smoking, and herbal supplements can inhibit or induce a drug's metabolism by CYP isozymes. This forms the basis of many major drug interactions. This will be discussed in the Drug Interactions and Pharmacogenetics chapters.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

CYPs are not the only drug-metabolizing enzymes, although they are probably the most important in terms of the percentage of drugs they metabolize. Butyrylcholinesterase (BChE), an enzyme related to acetylcholinesterase whose function is metabolizing the neurotransmitter acetylcholine, is a plasma enzyme responsible for metabolizing drugs with esterase bonds, such as succinylcholine, mivacurium, pilocarpine, cocaine, and novocaine. Individuals deficient in BChE don't become aware of the deficiency until usually after administering succinylcholine, a muscle relaxant used in surgery. Recovery is typically rapid, but patients with BChE deficiency take longer to recover and may require mechanical ventilation before regaining consciousness.

UDP-glucuronyltransferase (UGT) is a cytosolic enzyme important for Phase II conjugation reactions, transferring glucuronic acid to the oxygen, nitrogen, sulfur, or carbonyl groups of a drug. Morphine was used as an example of this reaction earlier in this section. Physiologically, UGT is an important detoxification mechanism for bilirubin, a waste product of heme metabolism. In humans who are deficient in UGT, a jaundice-like condition known as Gilbert's Syndrome can develop. A more serious, potentially fatal, autosomal recessive genetic disorder known as Crigler-Najjar syndrome may develop when individuals are completely deficient in UDGPT. UDGPT is highly polymorphic, with more than 200 different alleles being reported for the UGT1 and UGT2 gene families.

Lastly, N-acetyltransferase (NAT) is a Phase II cytosolic enzyme that transfers an acetyl group to drugs and exists in two forms, NAT-1 and NAT-2. Physiologically, NAT contributes to the metabolism of serotonin to form melatonin and may also play a role in folate metabolism. NAT metabolizes several carcinogens in cigarette smoke and car exhaust fumes. NAT-1 and NAT-2 have overlapping substrate specificity, but the major difference is that NAT-1 is found early in fetal tissue, whereas NAT-2 is not observed until after the first year of life. Human NAT-2 is the major isozyme responsible for metabolizing many drugs, including procainamide, sulphonamide antibiotics, isoniazid, hydralazine, nitrazepam, and dapsone. NAT-1 and NAT-2 are polymorphic, exhibiting two phenotypes: rapid acetylators and slow acetylators. Approximately 50% of Caucasians and African Americans are slow acetylators, whereas most Eskimos and Asians are rapid acetylators. NAT polymorphism has a significant consequence, as an increased risk of certain forms of cancer has been observed in slow acetylators.

The discovery of genetic polymorphisms was largely a matter of luck. In 1975, scientists of Saint Mary's Hospital Medical School in London ingested 40 mg of debrisoquine, an antihypertensive. Bob Smith, the lab director, became dizzy, faint, and hypotensive, which lasted for days. When they examined everyone's urine, they found that Bob Smith's urine had unusually low concentrations of the 4-hydroxydebrisoquine metabolite compared to the others'. Further research identified two subpopulations: a large group that could "extensively" metabolize debrisoquine and a smaller group that could not.

Shortly after that, Michel Eichelbaum at the University of Bonn made a similar discovery with sparteine, in which some subjects complained of blurred vision, double vision, dizziness, and headaches. When plasma sparteine concentrations were measured, those subjects had concentrations 5 times higher than those of others. Further research has shown that these polymorphisms result from changes in CYP2D6 alleles that affect the metabolism of CYP2D6 substrates [121].

Image courtesy of Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

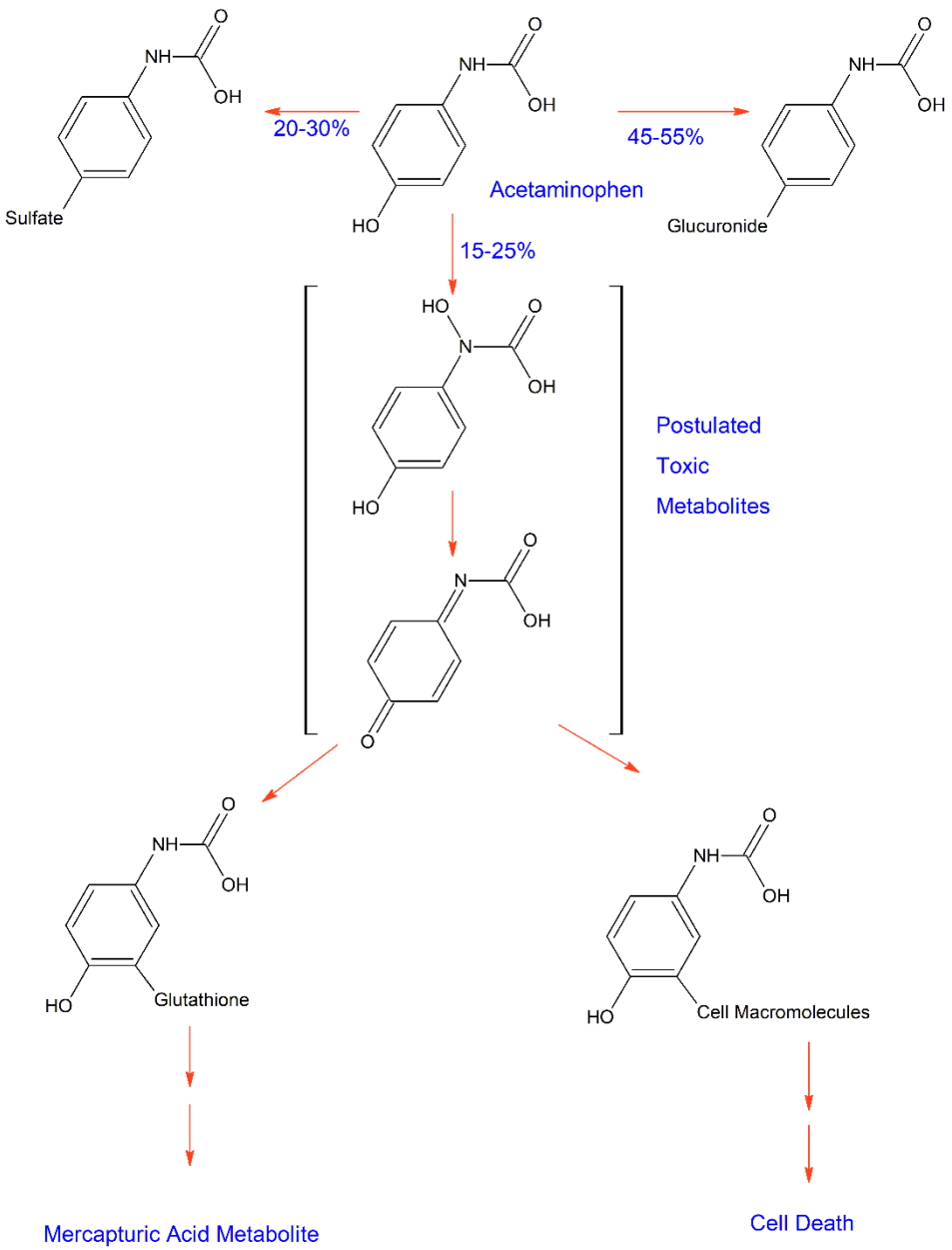
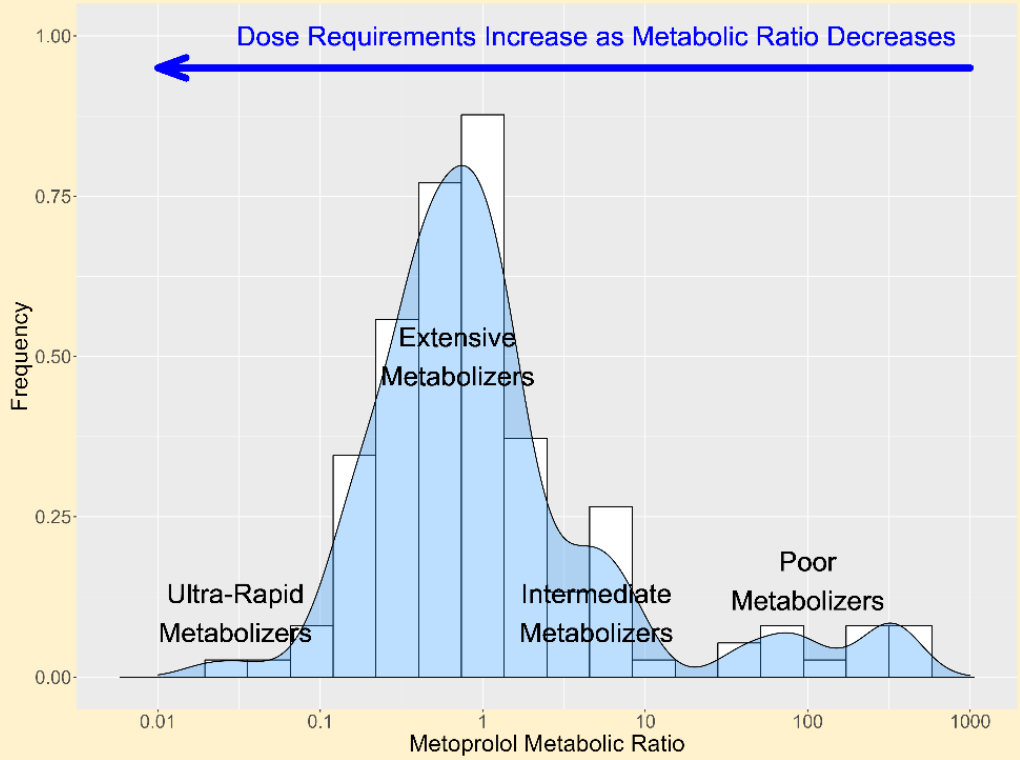


Figure 29. Metabolism of acetaminophen. Acetaminophen is metabolized to glucuronide and sulfate Phase II drug conjugates, which are normally eliminated in the urine. But there is only so much glutathione and sulfate in the body. When these pools are depleted, metabolism shifts to form reactive metabolites that can react with proteins in the liver, forming adducts and causing hepatic failure.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Metabolic Ratio for Debrisoquine in 143 Caucasians



People metabolize drugs at different rates. One reason is genetic polymorphisms, which are genetic differences in DNA that produce different enzyme variants with different affinities and/or capacities to metabolize their substrates. A single 100 mg dose of metoprolol tartrate was administered to 143 Caucasians. Their urine was collected over 8 hours. The urinary concentration of metoprolol (parent) and 4-hydroxy metoprolol, the major metabolite of metoprolol, was measured. The ratio of parent concentration to metabolite concentration, called the metabolic ratio or metabolite ratio, was calculated and plotted in the histogram above. The smaller the ratio, the more metabolites are being formed. Four groups of subjects become apparent:

- Ultrarapid metabolizers, which make up about 10% of the population, rapidly metabolize metoprolol, forming large amounts of metabolite and having small metabolic ratios;
- The predominant groups were extensive and intermediate metabolizers, which collectively form about 80% of subjects; and
- Poor metabolizers form little metabolite and have large metabolic ratios.

Generally, the therapeutic dose for a patient increases with an increasing metabolic ratio. Hence, ultrarapid metabolizers need larger doses than extensive and intermediate metabolizers to maintain therapeutic effects. Poor metabolizers require smaller doses, as they do not metabolize the drug appreciably. Data reported from [122].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Although metabolism is crucial for detoxifying potentially hazardous foods and drugs, it can have a dark side – some metabolites are toxic or can become toxic under certain circumstances. Toxicity from metabolites occurs by multiple means:

- During the metabolic process, an intermediary product is formed that is highly reactive (technically, they are either free radicals or electrophiles) and can attack cellular macromolecules like proteins, nucleic acids, or unsaturated lipids.
- A metabolite can accumulate in cells to such a degree that it reacts with cellular components before being metabolized it to a less toxic, non-reactive metabolite.
- A metabolite not further metabolized can accumulate in cells and react with cellular components.

When the metabolite binds to the protein, the enzyme's activity changes, often decreasing its function. Examples of drugs that form toxic metabolites include halothane and valproic acid, but special attention needs to be given to acetaminophen.

Acetaminophen Facts

- There are more than 75,000 visits to emergency rooms in the United States every year from acetaminophen overdose.
- Of these, 2600 patients will require hospitalization.
- Of these, 100 patients die every year from acetaminophen-induced liver failure.
- The typical dose for acetaminophen is 325 to 650 mg every 4 to 6 hours or 1000 mg every 6 to 8 hours. An overdose is considered to be more than 7000 mg a day, although the smallest recorded dose resulting in hepatotoxicity has been 5400 mg. Hence, there is not a large difference between a safe dose and an overdose.

Every year, more than 75,000 people go to the emergency room for an acetaminophen overdose, making it the leading cause of acute liver failure in the United States. The cause of such drug-induced liver failure is due to the metabolism of acetaminophen. Acetaminophen is normally metabolized by Phase II conjugation to sulfate and glucuronide metabolites in the liver. A small percent of the dose is metabolized by a Phase II reaction forming a nontoxic glutathione metabolite excreted in the urine (see Figure 29).

There is only a limited amount of sulfate and glutathione available in the liver for drug metabolism. When a sufficiently high dose is administered, the sulfate and glucuronidation pathways become saturated, shifting metabolism toward the glutathione metabolite and ultimately depleting glutathione stores. After glutathione pools are depleted, intermediate metabolites form reactive species that attack mitochondrial proteins in the liver, causing hepatotoxicity and, in some cases, death. Because the ratio of the dose that leads to serious side effects to the therapeutic dose is small, acetaminophen is called a 'narrow therapeutic index' drug. Treatment involves removing the remaining acetaminophen from the gastrointestinal tract by administering activated charcoal, which binds to the drug and facilitates its elimination through the feces, and by administering N-acetylcysteine, which replenishes the body's glutathione stores.

Prodrugs are inactive drugs that become active when metabolized.

About 5-7% of all approved small-molecule drugs are prodrugs. Examples include codeine, L-DOPA, terfenadine, and many cancer and HIV drugs, including 5-fluorouracil and zidovudine.

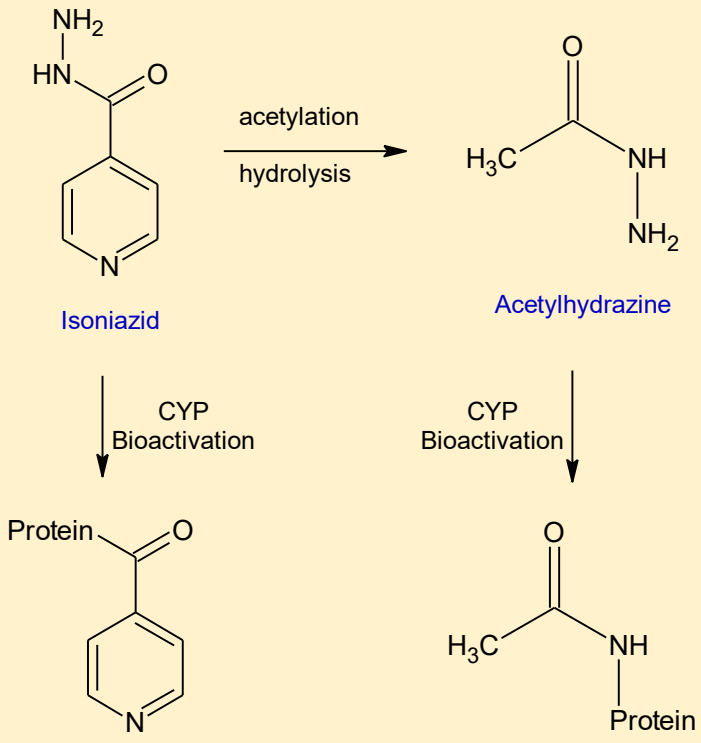
While metabolism can lead to negative consequences, it can also be leveraged to enhance a drug's therapeutic or pharmacokinetic profile. Metabolism can be used to design drugs that are difficult to absorb orally or that require pharmacological activity in specific organs. These latter drugs, known as prodrugs, are pharmacologically inactive but are metabolized into an active metabolite [123]. One example is codeine, a commonly prescribed drug for pain, diarrhea, or as a cough suppressant, which is itself inactive but, in the liver, is metabolized by CYP2D6 to the pharmacologically active metabolite morphine and by glucuronidation to the active metabolite 6-glucuronide-codeine (which may be the major metabolite responsible for activity). L-

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

DOPA, which is used to treat Parkinson's disease, is another example of a drug that is metabolized to an active metabolite at the site of action. DOPA itself is inactive but is metabolized by DOPA decarboxylase to dopamine, the active form, a neurotransmitter in the brain necessary for proper motor movement. So why not administer dopamine directly? Dopamine does not cross the blood-brain barrier, but L-DOPA does. Hence, L-DOPA crosses the blood-brain barrier and enters the brain, where it is metabolized into dopamine, thereby eliciting its effect.

Drug-Induced Liver Injury (DILI) with Isoniazid

Isoniazid was discovered in 1912, and its activity against tuberculosis was first reported in the 1950s. By the 1970s, hundreds of cases of liver injury were reported after long-term isoniazid use. The clinical syndrome of idiosyncratic DILI includes fever, malaise, nausea, and vomiting. In many cases, liver injury is asymptomatic, only detected by increases in the liver enzymes aspartate aminotransferase and alanine aminotransferase. Elevations in liver enzymes can occur as early as one week and as late as nine months after starting treatment. In most cases, liver injury resolves without sequelae after treatment is discontinued. But, in 1% of patients, liver injury is severe, life-threatening, and may require a liver transplant. The cause of the change from mild to severe phenotype is unknown, although the mechanisms for liver injury are generally attributed to the covalent binding of metabolites to liver proteins.

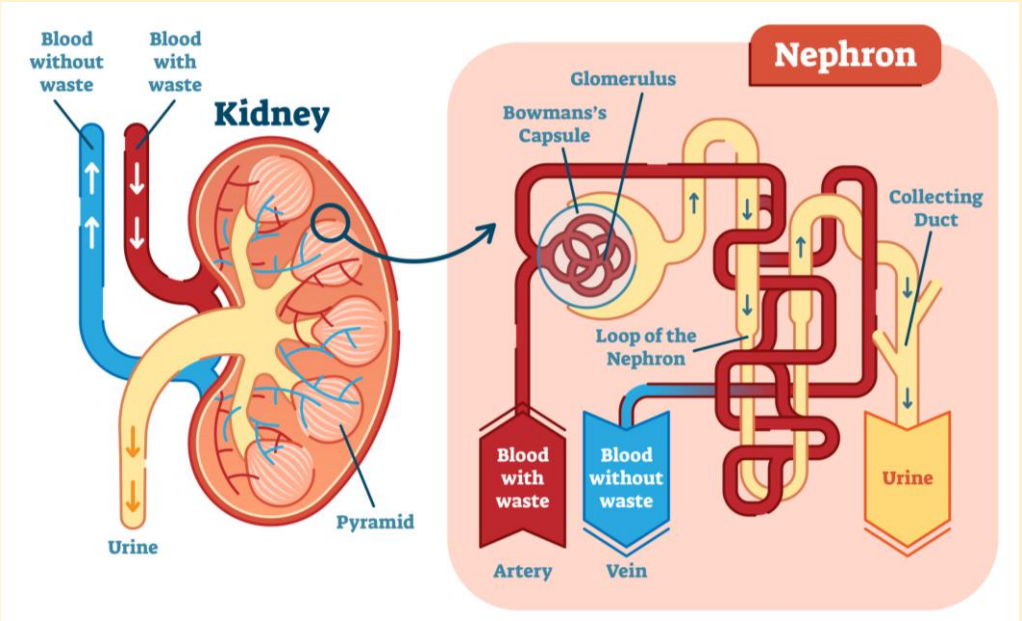


Three metabolites of isoniazid are proposed: hydrazine, acetyl-hydrazine, and a metabolite from direct bioactivation of isoniazid itself. Once formed, the free amine groups can covalently bind nonspecifically to lysine residues in liver proteins. The degree of binding directly correlates to the degree of liver injury. Isoniazid N-acetyltransferase (NAT) slow metabolizers appear to have a higher incidence of liver injury due to higher acetylhydrazine concentrations. Furthermore, treatment with drugs that induce CYP enzymes also increases the risk of liver injury.

In severe liver injury cases, there also appears to be an immune-mediated component to the syndrome. It is believed that isoniazid-protein adducts become immunogenic in some patients, leading to an autoimmune reaction. Evidence for this comes from a lupus-like reaction observed in some patients, the detection of auto-antibodies to the protein adducts in others, and rechallenge following cessation of the drug can result in symptoms reappearing hours after the drug is taken.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Diagram Showing the Kidney and Basic Unit of the Kidney – the Nephron



The basic unit of the kidney is the nephron, of which there are about 1 million per kidney. The nephron filters blood to form urine, regulating fluid balance, electrolytes, acid–base status, and waste excretion to maintain overall homeostasis. Drugs can also be excreted in urine. Drugs with molecular weights < 60,000 Da can be filtered through the glomerulus, where they can be reabsorbed back into the blood via tubular reabsorption or excreted into urine. Drugs in arterial blood feeding the kidney can be actively extracted from the blood into the tubular lumen and excreted into urine (a process called active secretion). The kidney can also metabolize drugs by Cytochrome P450 enzymes.

Another interesting prodrug is terfenadine, marketed in the 1980s and 1990s as Seldane for the treatment of allergic rhinitis. It was one of the first billion-dollar-a-year drugs because it did not cross the blood-brain barrier and cause sedation, unlike other antihistamines like Benadryl (diphenhydramine). After it was approved, case reports began to appear of patients treated with Seldane dying from cardiac-related disturbances. Afterward, the manufacturer, Hoechst Marion Roussel, began investigating the mechanisms involved and discovered that terfenadine inhibited the inward delayed rectifier potassium channel (Ikr) in the heart. When co-administered with certain drugs that inhibited its metabolism, like erythromycin, terfenadine concentrations could increase to such levels that blocked slow cardiac repolarization and, in some cases, cause an often-fatal cardiac arrhythmia called Torsades de Pointes. Fortunately for the manufacturer, terfenadine itself was inactive (it was a prodrug), and all pharmacologic activity was due to its primary metabolite, fexofenadine. Fexofenadine did not affect the I_{kr} channel and had no cardiac liabilities. So, Hoechst Marion Roussel submitted a New Drug Application for fexofenadine, which was approved and marketed as Allegra. Shortly after that, Seldane was removed from the market. Today, Allegra is an over-the-counter drug used by millions worldwide.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drug Elimination

The major organs of elimination are the liver and kidneys, although some drugs are partially excreted into saliva, expired air, breast milk, and sweat.

The Kidney. Everyone knows the kidney "cleans the blood." The kidney eliminates drugs through many processes: glomerular filtration, tubular secretion, tubular reabsorption, and metabolism. Renal elimination is the sum of these processes:

Elimination = filtration + secretion + metabolism - reabsorption

The basic physiologic unit of the kidney is the nephron, which performs all the kidney's functional activities. In total, there are over a million nephrons per kidney. In each nephron, blood enters the kidney via the renal artery, is distributed into smaller arterioles, and then enters the glomerulus, where it is filtered, and the process of urine formation begins. The filtrate then enters the proximal tubule, the loop of Henle, the distal tubule, and the collecting ducts, and exits as urine, which is stored in the bladder until urination. As the glomerular filtrate moves through the nephron, it becomes increasingly concentrated by tubular reabsorption of water until it becomes urine, the end product.

For drugs that undergo tubular reabsorption in the nephron, the pH of the filtrate can be clinically manipulated by administering buffer solutions, like sodium bicarbonate, for therapeutic benefit. This change in filtrate pH can change the percentage of the drug in the ionized form and the amount of drug available for reabsorption.

For example, salicylate poisoning may occur in suicide attempts or by accidental ingestion of overdoses of aspirin. Aspirin is a weak acid and has a pKa of 3.5. Normal urine pH is about 6.5 to 7, and 99.9% of aspirin exists in an ionized form that does not undergo tubular reabsorption. But with severe salicylate poisoning, the pH can be less than 6, a large percentage of the drug exists in a nonionized form, and significant tubular reabsorption can occur, thereby prolonging the toxicity. As part of the standard of care for salicylate overdose, physicians administer sodium bicarbonate to alkalinize the urine pH from 7.5 to 8.5, thereby increasing the percentage of salicylate in the nonionized form to essentially 100%, preventing reabsorption.

At the first step, the glomerulus filters out drugs with a molecular weight of more than 60,000 Daltons, preventing them from entering the nephrons. Large macromolecules, such as albumin or antibodies, do not pass through the glomerulus but remain in the blood and are carried away from the kidneys by the renal vein. What passes through the glomerulus has a composition similar to that of plasma water and enters the proximal tubule, which is also the site of tubular secretion. This active process transports drugs, mostly weak electrolytes and acids, from the blood to the tubule against a concentration gradient. Approximately two-thirds of the water in the filtrate is reabsorbed in the proximal tubule. In the Loop of Henle, 25% of water from the proximal tubule is reabsorbed as it enters the distal tubule, where a further 8% is reabsorbed, along with the reabsorption of some drugs. What remains is then concentrated into the collecting ducts, where the final 2% of water is reabsorbed. After the collecting ducts, urine is collected in the bladder. If it sounds as though a lot of water is reabsorbed as it moves through the kidney, that is correct. Consider this: the kidneys receive 20% of the cardiac output. Over the course of a day, the kidney filters 180 L of plasma, of which 1 to 2 L is excreted as urine. This means more than 99% of the plasma flowing through the kidney is reabsorbed daily.

The major factors influencing renal elimination are molecular size, protein binding, pH, pKa, and lipophilicity. As mentioned, large macromolecules cannot cross the glomerulus and have little renal excretion. Highly protein-bound drugs cannot cross the glomerulus and have low renal elimination because they remain bound to proteins. Tubular reabsorption from the nephron back into the blood is largely a passive process, although some active transporters, such as organic anion transporter-4 (OAT-4), have been identified. A drug's pKa is the pH at which 50% of the drug is ionized. Hence, pH and pKa play a large role in the passive tubular reabsorption of drugs across renal epithelial cells.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

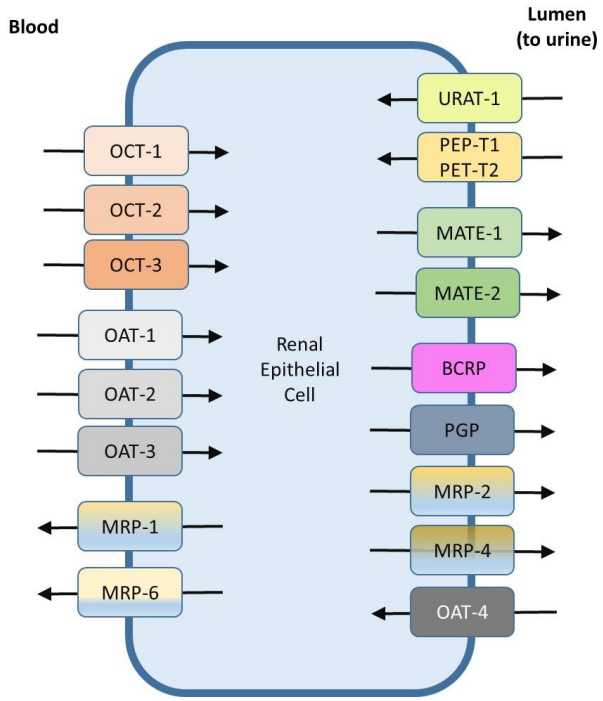


Figure 30: Major kidney transporters involved in tubular reabsorption (from lumen to blood) and tubular secretion (from blood to lumen and, ultimately, the urine). Transporters of the same family are coded using different shades of the same color. Influx transporters pump drugs into cells. Efflux transporters pump drugs out of cells.

If the drug exists predominantly in an ionized form, it will not be reabsorbed (recall that only unionized drugs can cross cell membranes). Furthermore, the unionized form of the drug must have sufficient lipid solubility to cross the membrane; highly polar, nonionized drugs will not be reabsorbed. The larger the difference between the drug's pKa and the filtrate pH, the less likely the drug will have a significant tubular reabsorption component.

Tubular secretion occurs throughout the nephron, from the proximal tubule to the collecting duct, and involves the transfer of ions and drugs from the blood into the tubules, thereby bypassing the glomerular filtration process. Tubular secretion is generally an active process that involves transporter proteins carrying the drug from the blood to the filtrate against a concentration gradient (Figure 30). Transporters are proteins that span the cell membrane and transport small molecules, like drugs, from one side of the membrane to the other. They can transport drugs from outside the cell to inside the cell or vice versa. Several transporters have been identified in the last 20 years. The important ones are the organic acid transporter (OAT), distinct from the OATP transporter; the organic cation transporters (OCT); p-glycoprotein (P-gp); breast cancer resistance protein (BCRP); and multidrug and toxin extrusion (MATE) proteins.

Typical ions removed by tubular secretion include hydrogen, potassium, ammonium, creatinine, urea, and some hormones. Drugs that undergo tubular secretion include penicillin, beta-lactam antibiotics, probenecid, thiazide diuretics, morphine, and procaine. Kidney transporters can be manipulated for therapeutic benefit. Thiazide diuretics achieve their efficacy by inhibiting the Na^+/CL^- cotransporter in the distal tubule. This transporter is responsible for sodium reabsorption, and its inhibition results in reduced water reabsorption, leading to diuresis. Another example is penicillin, which undergoes tubular secretion via the OAT-3 transporter. Probenecid, which also undergoes tubular secretion with the OAT-3 transporter, competes with penicillin when co-administered. Hence, physicians should

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

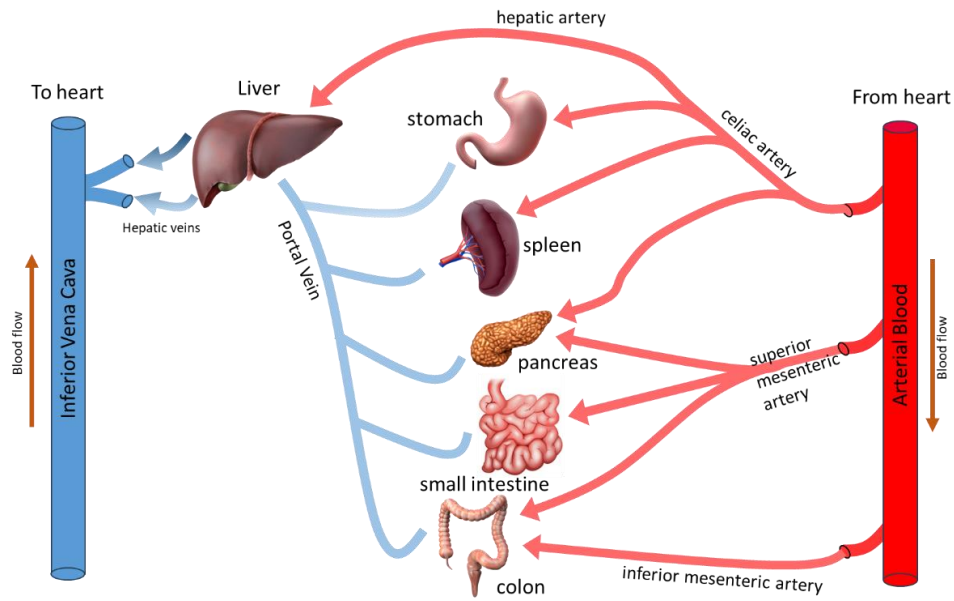
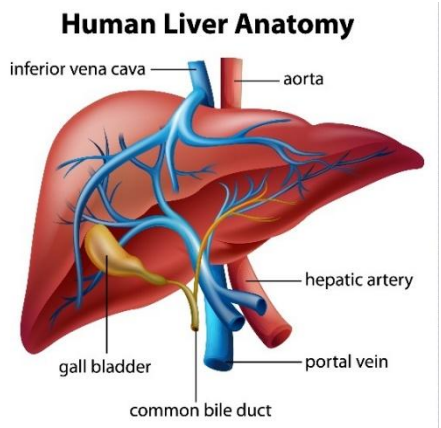


Figure 31: Blood flow to and from the liver. Blood flow to the liver is unique in that it is supplied by both the hepatic artery and the portal vein, which collects venous blood from the stomach, spleen, pancreas, small intestine, and colon. Because of blood in the portal vein that feeds the liver, the liver can metabolize orally administered drugs before they are distributed to the rest of the body via arterial blood; a process known as 'first-pass metabolism.'

prescribe both medications to "boost" penicillin's plasma concentration by 2- to 4-fold. Competition for these transporters by other drugs is a frequent source of drug interactions.

Drug metabolism can also occur in the kidneys, although it is a minor route of elimination for most drugs compared to the liver. Cytochrome P450, sulfatases, methyltransferases, glutathione S-transferase, and UDP-glucuronyl-transferases have all been identified in the kidney and can contribute to drug Phase I and II metabolism. Renal metabolism is often suspected when metabolites not occurring in plasma are identified in the urine.

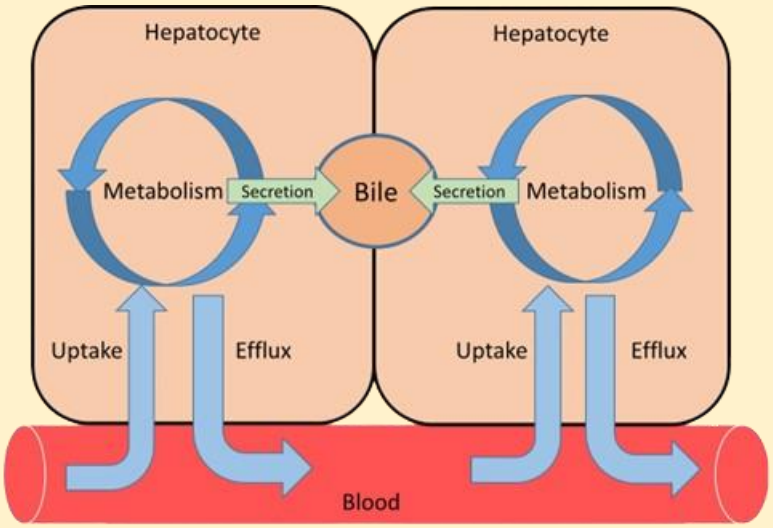


The Liver. The liver is located in the right upper quadrant of the abdomen. It is uniquely located because its blood flow comes from two sources: one is oxygenated blood via the hepatic artery and the other is by the hepatic portal vein, which carries blood from the intestines, gallbladder, pancreas, and spleen, and contains blood with nutrients and toxins after digestion. The hepatic portal vein allows the liver to serve as a protective barrier against drugs and toxins. At any given time, about 25% of blood from the heart goes to the liver, and about 10% of blood is actually inside the liver (Figure 31).

The liver has many functions, but it is also the major site of drug metabolism in the body and is responsible for the first-pass metabolism of orally administered drugs. Beyond this, however, the liver also eliminates drugs by biliary secretion. Bile is made in the liver and secreted in the gall bladder. After a meal, bile is secreted into the small intestine to aid in the digestion of fats and lipids. Drugs that are neither polar nor lipid, called amphipathic molecules, and have a molecular weight (MW) in the 500 to 800 Da range, are candidates for biliary secretion. Such drugs include montelukast (MW 586 Da),

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

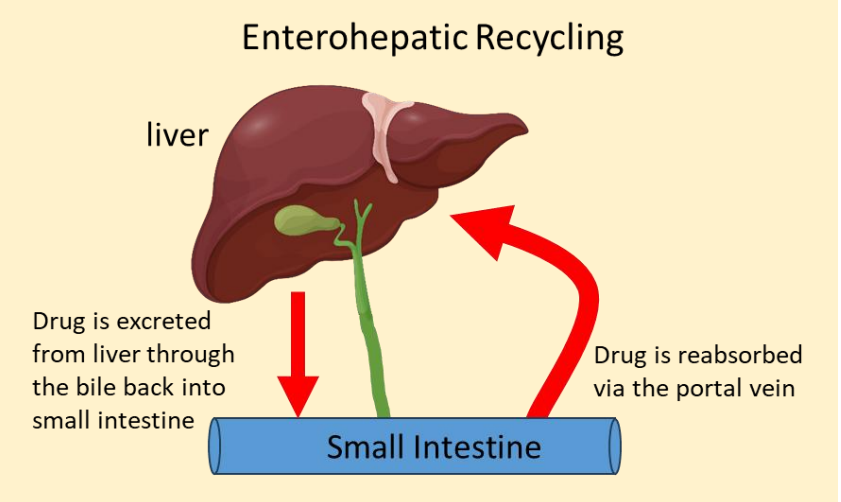
Events That Contribute to Hepatic Drug Elimination



Drugs in the hepatic artery or portal vein can enter liver cells (hepatocytes) and be metabolized or secreted into the bile, a process called biliary secretion. Both processes act to remove the drug from the blood,

fexofenadine (502 Da), cyclosporine (1202 Da), and erythromycin (734 Da). Conjugated drug metabolites, particularly glucuronide and glutathione conjugates, are also subject to biliary excretion; these include indomethacin, dapsone, carbamazepine, and chloramphenicol. Once in the bile and secreted into the small intestine, these drugs and metabolites may be reabsorbed or excreted into the feces. Many of the same transporters that play a role in renal excretion are responsible for biliary secretion. These are mostly the OATs, OCTs, MRP, P-gp, and BCRP. The OATs, OCTs, MRP1, MRP3, and MRP6 are primarily responsible for transport from the hepatic sinusoid (small blood vessels in the liver) into hepatic cells, while P-gp, MRP2, and BCRP secrete the drug from inside the cell into the bile.

For several drugs that undergo biliary secretion, a few undergo an interesting phenomenon called enterohepatic recycling, whereby the secreted drugs are reabsorbed back into the body in the small intestine, creating a cycle of absorption, secretion, and reabsorption. Specific natural products that undergo enterohepatic recycling include steroid hormones, thyroid hormones, insulin growth factors, folate, and vitamin D3. Drugs that undergo enterohepatic recirculation include ezetimibe, irinotecan, possibly morphine, and possibly ranitidine. Although some drugs undergo



Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

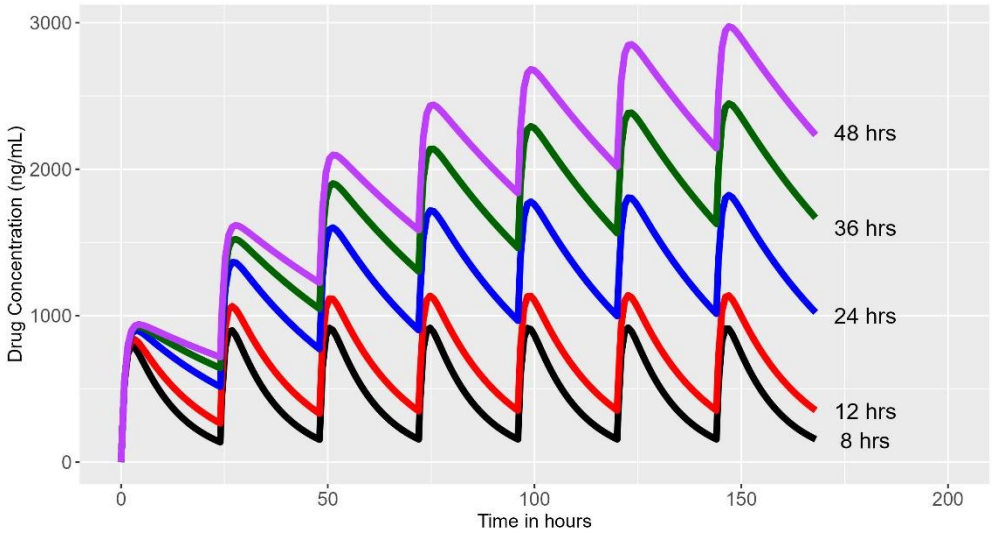


Figure 32. Concentration-time profiles for a hypothetical drug with varying degrees of half-life. A total of 100 mg of the drug was assumed to be taken once daily every 24 hours. The absorption rate constant was 1 per hour, and the volume of distribution was 100 L. Half-life was varied from 8 to 48 hours.

enterohepatic recycling directly, most are metabolized to Phase II conjugates, particularly glucuronides, and then secreted into bile. Normal gut flora produces an enzyme called β-glucuronidase that cleaves the glucuronide metabolite back to the parent form, which is then reabsorbed in the small intestine. The net effect of this phenomenon is to prolong the drug's retention in the body.

Some drugs inhibit enterohepatic recycling, which can lead to positive and negative effects. Cholestyramine, a bile acid sequestrant used to treat hypercholesterolemia, inhibits the recycling of digitoxin and is used in cases of overdose to reduce toxicity. In contrast, doxycycline, a broad-spectrum antibiotic, can cause contraceptive failure in women who take contraceptives containing estrogen. Normally, estrogens undergo enterohepatic recycling after glucuronide conjugation, followed by gut flora-mediated metabolism of the glucuronide conjugate. But doxycycline kills gut flora, breaking the cycle and preventing estrogens from being reabsorbed into the body. Hence, doxycycline prevents their contraceptive effect and can cause a breakthrough pregnancy.

Kinetics of Multiple Dosing and the Concept of Steady-State

Most drugs are not taken in a single dose. One exception is painkillers, where you might take a single dose of acetaminophen or ibuprofen to relieve a headache or some other ache or pain. Instead, most drugs are taken repeatedly, usually once or twice a day, for some extended period. Earlier in this chapter, the kinetics after single-dose administration were presented. How do the concentrations of a drug change upon repeated dosing?

With repeated administration, most drugs accumulate in the body, and the concentration-time profiles begin to superimpose on each other (Figure 32), until a steady-state is reached. The concentration-time profile after each successive dose is the sum of the concentration-time profile after a single dose and the amount remaining after previous doses – this is called the principle of superposition. The greater the half-life relative to the dosing interval τ, the greater the degree of accumulation. Eventually, the amount of drug taken each day equals the amount eliminated each day. At that point, the drug is said to be in steady-state. Technically, at steady-state, the amount in

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Pharmacokinetics and the Death of Michael Jackson

On 25 June 2009, the entertainer Michael Jackson was found dead in his home in Los Angeles. His physician, Conrad Murray, said he found Jackson not breathing and administered CPR, but to no effect. Jackson was taken to a hospital, where he was declared dead. An autopsy soon revealed that Jackson died of an overdose of the anesthetic propofol and the anxiolytic lorazepam. His death was declared a homicide. The investigation soon focused on Conrad Murray, who was quickly charged with involuntary manslaughter. During the trial, prosecutors claimed that Murray administered the drugs himself and accidentally overdosed Jackson. The defense claimed that Jackson suffered from insomnia and self-administered the drugs, leading to his death. Population pharmacokinetics was used to confirm the defense's theory. Steven Shafer, an anesthesiologist and expert in clinical pharmacology and pharmacokinetics, testified for the prosecution and, using pharmacokinetic modeling and simulation, examined whether the defense's claim was consistent with autopsy results. Shafer was unable to confirm that Jackson's self-administration could produce the propofol and lorazepam blood concentrations measured at autopsy. The only scenario that was consistent with the data was one in which Conrad Murray continued to administer propofol after Jackson was sedated. On 11 November 2011, Murray was found guilty and sentenced to 4 years' imprisonment.

equals the amount out.

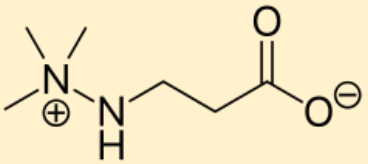
If you've ever sat in a bathtub and run water into the tub while the water was leaving through the drain, when the volume of water running into the tub is the same as the volume of water leaving the tub, and there is no net loss or gain of water, that is the point of steady-state. The same principle applies to drug accumulation with multiple dosing. In Figure 32, steady-state is when the concentration-time profiles after each administration appear the same after repeated administration. For instance, look at the profile for a drug having a half-life of 24 hours (the blue line). The last three doses appear to be similar – they have the same Cmax and the same concentration immediately before the next dose (referred to as the trough concentration or C_{trough}). Hence, steady-state appears to have been reached by 100 hours.

What might not be immediately obvious is that the time to steady-state is dependent solely on a drug's half-life. After the first dose, 50% of the steady-state has been reached. After two doses, 75% of the steady-state has been reached. After three doses, 87.5% of steady-state has been reached, etc. By the time five half-lives have been reached, 98% of steady-state has been reached. Hence, steady-state is assumed to be reached in 4 to 5 half-lives. The longer the half-life, the longer it takes to reach steady-state. In Figure 32, the drug having a half-life of 48 hours (purple line) never appears to reach a steady-state. This is consistent because it would take ~200 hours to reach steady-state, and the dosing stopped at 168 hours in the figure.

The half-life of the drug is often the defining reason for how often a drug is dosed clinically. If the drug's half-life is less than the dosing interval (see the 8-hour half-life concentration-time profile (black line) in Figure 32), accumulation will be small to none, and with multiple doses, each dose will look like a series of single doses. In this instance, the drug will need to be dosed two or three times a day for any significant accumulation to occur. For those drugs whose half-life is longer than the dosing interval, accumulation will occur, and concentrations will be higher after multiple dosing compared to single dosing. For these drugs, the dosing interval is usually daily for an oral drug. And for drugs whose half-life is very long, such as with monoclonal antibodies, doses may be given weekly or every few weeks.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Maria Sharapova Fails Drug Test at Australian Open



Meldonium, also known as mildronate, is a potential performance-enhancing drug that improves blood flow and exercise capacity in athletes. The World Anti-Doping Agency (WADA), responsible for harmonizing and coordinating anti-doping rules and policies across all sports and countries, found "evidence of its use by athletes with the intention of enhancing performance" by carrying more oxygen to the muscles. The decision to ban meldonium was made on 16 September 2015 and went into effect on 1 January 2016.

The tennis professional Maria Sharapova was set to compete at the Australian Open in March 2016 and, in January, submitted a urine sample for testing. She failed and tested positive for meldonium. On 7 March 2016, it was reported by the press that she failed her drug test. On 12 March 2016, she was provisionally suspended from competing pending the outcome of the case.

Could Sharapova have used pharmacokinetics to acquit herself? Suppose she last took the drug on New Year's Eve, the day before the ban went into effect. If this were the case, then she did nothing wrong. The drug in her urine was leftover meldonium from the last time she "legally" took the drug.

Could this be true? Meldonium's half-life is 4 to 6 hours. According to reports, she took her drug test 26 or 27 days, or 104 to 156 half-lives, after the ban went into effect. This would be an impossible claim to make because 0.00000000000000000000000000078886% of the drug remains in the body after 100 half-lives. But…. While the half-life of meldonium appears to be four to six hours, Grindeks, the Latvian manufacturer of the drug, claimed that meldonium may linger in the body for months. Is that possible? Well, technically, yes, all drugs may last in the body forever, but their concentrations may be so low as to be close enough to zero. In Sharapova's case, if she took 500 mg twice daily, then $\sim 8 \times 10^{-27}$ mg, not zero mg, would have remained in her body when she took the drug test. Certainly, it is nowhere near enough to trigger a positive drug test.

How was this case resolved? Sharapova admitted she continued taking meldonium after the ban date went into effect, but claimed she did not know it was on the banned list and did so for a medical condition, stating that she had been taking the drug since 2006. The International Tennis Federation ultimately rejected her claims, and she was banned from tennis for two years. She returned after the ban and played for two more years until 2020, when she retired from tennis. Today, this story is largely forgotten, but for a brief period in 2016, pharmacokinetics was in the national news!

Figures courtesy of Wikipedia Creative Commons.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

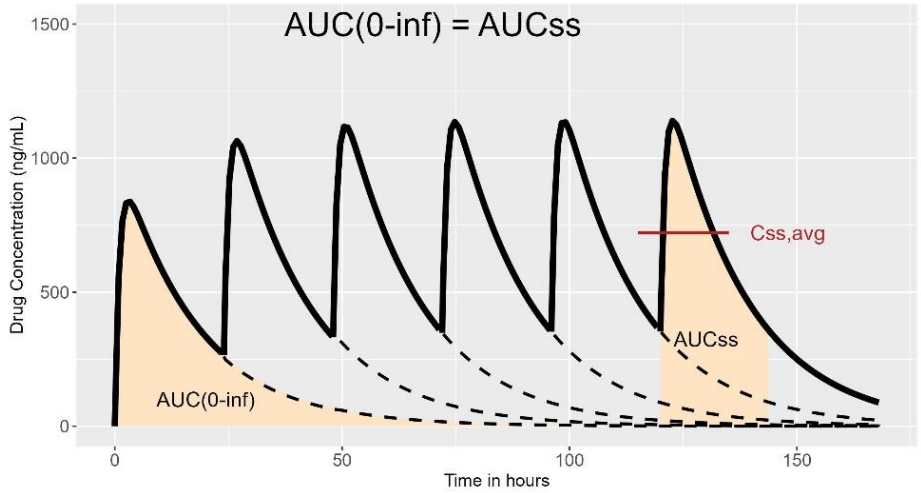


Figure 33. Concentration-time profile after multiple doses of a drug given once daily. For a drug having linear pharmacokinetics, the AUC of the concentration-time profile from time 0 to infinity (shown as the beige band) is equal to the AUC under the dosing interval at steady-state (AUC_{ss}, shown in the beige band after the 6th dose). For this simulation, a dose of 100 mg was given once daily. The volume of distribution of the drug was 100 L, the absorption rate constant was 1 per h, and the half-life was 12 h. The average steady-state concentration ($C_{ss,avg}$) is represented by the red line.

The degree of accumulation, defined as the accumulation ratio (AR), is the ratio of drug concentration at steady-state to that on Day 1. Mathematically, the ratio is equal to

$$AR = \frac{1}{1-\exp\left(\frac{-Ln(2)}{\text{half-life}}\times\tau\right)}. \tag{13}$$

If the half-life equals the dosing interval (τ), the accumulation ratio equals 2, meaning that concentrations at steady-state are twice as high after single dose administration (see the blue line in Figure 32 as an example). The average concentration at steady-state ($C_{ss,avg}$) is equal to

$$C_{ss,avg} = \frac{F\times D}{CL\times\tau} \tag{14}$$

where F is bioavailability, CL is clearance, and D is dose. If you notice, Eq.(14) can be simplified to

$$C_{ss,avg} = \frac{AUC}{\tau}. \tag{15}$$

Hence, the average concentration at steady-state can be considered a weighted time average of the AUC. Because this is a simple relationship, changing the dose and dosing interval by the same amount will result in the same $C_{ss,avg}$. For example, changing the dose from 100 to 200 mg and the dosing interval from 12 to 24 hours will result in the same $C_{ss,avg}$.

One last interesting property of multiple dosing is shown in Figure 33. The AUCinf of the concentration-time profile after a single dose is equal to the AUC of the concentration-time profile under the dosing interval after the same dose at steady-state. One way to test whether a drug has linear pharmacokinetics is to compare the two AUC values. If they are the same, the drug has linear pharmacokinetics. If they differ, the drug exhibits nonlinear pharmacokinetics.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Linear and Nonlinear Pharmacokinetics

For a drug with linear pharmacokinetics, doubling the dose will also double the exposure. For example, if at a dose of 100 mg, C_{max} was 1 μg/mL and AUC was 10 μg · h/mL, then Cmax would be 2 μg/mL, and AUC would be 20 μg · h/mL at a dose of 200 mg. A drug has linear pharmacokinetics if no processes related to the drug's pharmacokinetics are saturable. This means no saturable protein binding, no saturable metabolism, and no saturable transport, among other things. For most drugs, linear pharmacokinetics is the rule. However, many processes involved in pharmacokinetics utilize proteins, including enzymes, binding proteins, and transporters. As such, these proteins may be saturable at their binding or active sites, i.e., only a limited amount of protein is available to interact with the drug. When saturation occurs, nonlinear pharmacokinetics often follow. With nonlinear pharmacokinetics, you will get higher exposure than expected if you double the dose.

For the minority of drugs with nonlinear pharmacokinetics, a variety of mechanisms have been identified. For example, riboflavin undergoes saturable transport in the gut wall; propranolol undergoes saturable hepatic metabolism; phenylbutazone has saturable protein binding; methotrexate has saturable transport into and out of tissues; and penicillin undergoes saturable tubular secretion in the kidney. An entirely different nonlinearity mechanism is not due to protein saturation, but rather to changes in protein concentration. For example, carbamazepine increases the amount of enzymes involved in its own metabolism, a process known as induction, leading to higher metabolic clearance and a reduced half-life. Lastly, drug metabolites or other drugs may cause nonlinear pharmacokinetics by competing with the drug for saturable binding sites on an enzyme.

Drugs showing nonlinear pharmacokinetics have the following properties:

- AUC and C_{max} are dose-dependent;
- Half-life changes with dose due to changes in clearance or volume of distribution; and
- At low doses, the drug may show linear pharmacokinetics.

These properties are demonstrated when two drugs are shown side by side, one with linear pharmacokinetics and the other with nonlinear pharmacokinetics. The linear drug demonstrates dose-proportionality – the AUC increases by the same amount as the dose increases. The nonlinear drug, however, has a hockey-stick-shaped concentration-time profile: as the dose increases, the AUC increases disproportionately.

For linear drugs, their rate of elimination is equal to

$$\text{Rate of elimination} = CL \times C \,. \tag{16}$$

But for drugs that have nonlinear pharmacokinetics because of saturable metabolism, which is probably the predominant mechanism for nonlinearity, their elimination rate is equal to

$$\text{Rate of elimination} = \frac{V_{\max} \times C}{K_m + C} \,. \tag{17}$$

where V_{max} is the maximal rate of elimination and K_m, the Michaelis constant, is the concentration at which saturation starts to occur. Such drugs follow Michaelis-Menten kinetics.

Examples of drugs with nonlinear pharmacokinetics include alcohol, phenytoin, salicylate, and theophylline. It is interesting to look at the limiting cases in Eq. (17). When $C<<K_m$, Eq. (17) reduces to V_{max}/K_m, which is a constant. Hence, when drug concentrations are below Km, the drug's kinetics become linear. On the other hand, when $C>>K_m$, then the rate of elimination approximately equals V_{max}, which means that elimination becomes independent of concentration; in this instance, the elimination rate is said to be zero-order instead of first-order (Figure 34). The elimination rate depends on the drug concentration, as in the limiting cases.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

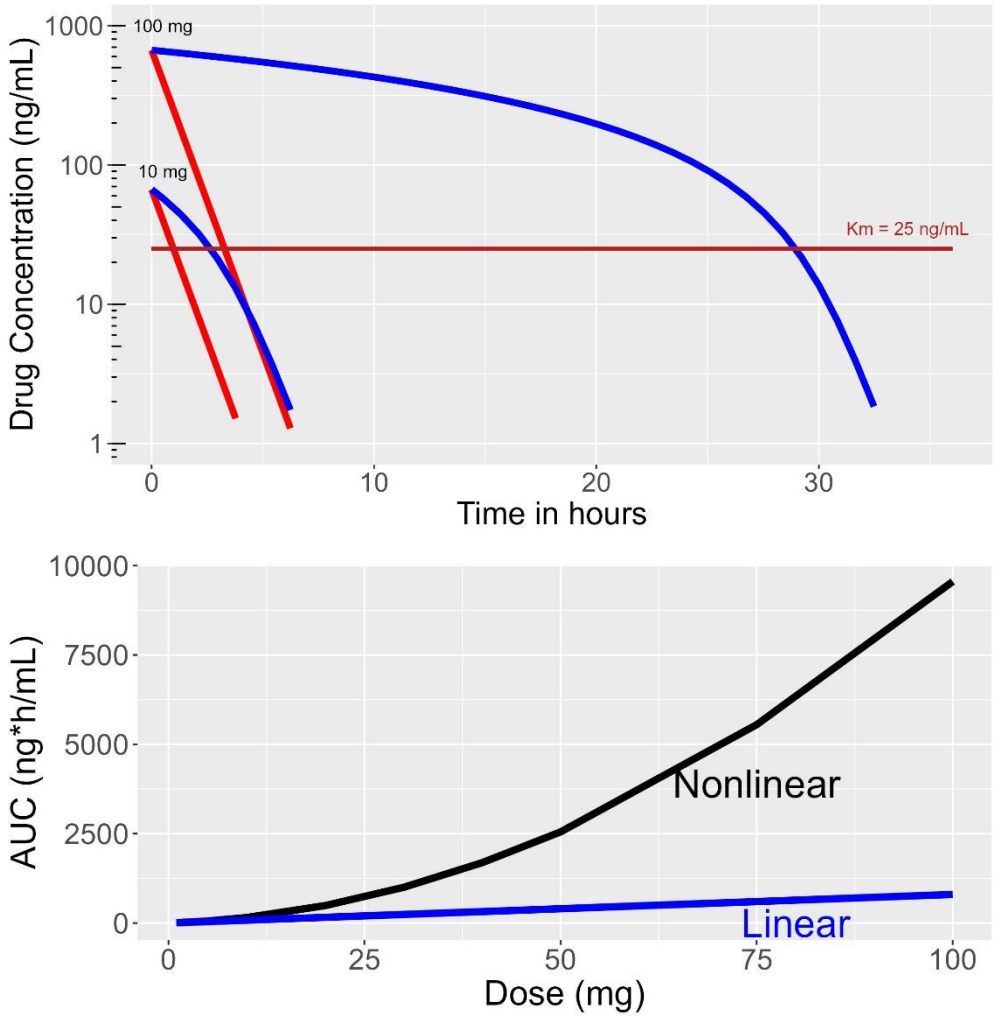


Figure 34: Plots demonstrating the difference between linear and nonlinear pharmacokinetics.

Top plot: Concentration-time profile for two drugs: linear pharmacokinetics (red lines) and nonlinear pharmacokinetics (blue lines) administered at 10 mg or 100 mg doses. The fire-red horizontal line shows the K_m, the concentration at which clearance is 50% of the maximal velocity, for the nonlinear drug. For the linear drug, concentrations decline linearly, and a 10-fold increase in dose from 10 to 100 mg results in a 10-fold increase in concentration. The same is not true for the nonlinear drug—a 10-fold increase in dose results in very drastic changes to the profile and a much longer elimination. Notice the hockey-stick-shaped profile for the nonlinear drug – this is a tell-tale sign of nonlinear pharmacokinetics.

Bottom plot: AUC as a function of dose for the same linear and nonlinear drugs. The nonlinear drug behaves linearly at very low doses, but as the dose increases, the AUC increases disproportionately.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

If you recall, the half-life of the drug can be described as $0.693 \times V_d/CL$. For nonlinear drugs

$$\text{half-life} = \frac{0.693 \times V_d \times (K_m + C)}{V_{\max}}. \quad (18)$$

Hence, the concept of half-life doesn't apply to nonlinear drugs because half-life becomes concentration-dependent.

There is also a class of drugs whose pharmacokinetics change over time; that is, their pharmacokinetics are not time-invariant. As such, the property that AUCinf equals the AUC over the dosing interval at steady-state (AUC_{ss}) is not true. These drugs are also considered to have nonlinear pharmacokinetics. Examples of these drugs include rifampin, which induces its own metabolism and increases its clearance, and omeprazole and clarithromycin, both of which inhibit their own metabolism.

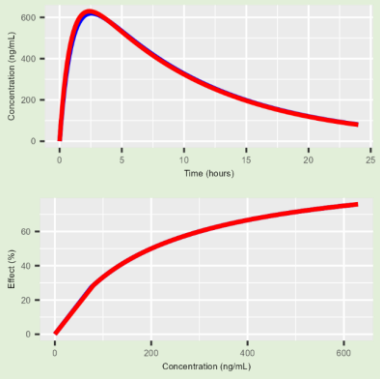
Drugs with nonlinear pharmacokinetics are challenging to dose in patients, and drug companies are reluctant to develop them because it is difficult to demonstrate efficacy without undue toxicity. Clinically, for nonlinear drugs with a narrow safety window, extra care must be taken when starting or changing a patient's dose. Furthermore, disease states can change a drug's V_{max} or K_m. For example, when carbamazepine is coadministered with phenytoin for the treatment of epilepsy, phenytoin Vmax increases due to carbamazepine-induced enzyme induction, potentially requiring a higher dose than expected. Drugs like cimetidine or chloramphenicol may increase phenytoin's K_m. Complicating things even further is that disease states, like hepatic cirrhosis, obesity, and even age, may change the V_{max} and K_m of a drug. For these reasons, many drugs are monitored for plasma or blood concentrations to ensure they remain within the safe and efficacious range. This is referred to as 'therapeutic drug monitoring'. An example is tacrolimus, which is used to prevent rejection of solid organ transplants. The therapeutic window for tacrolimus is 10 to 15 ng/mL. Patients' tacrolimus blood concentrations are routinely assessed to ensure they are within the recommended range. If they are not, their dose may be adjusted.

The Concept of Bioequivalence

Bioequivalence (BE) was introduced as a requirement for drug manufacturers to develop generic equivalents to brand-name drugs, a concept discussed in previous chapters. BE plays a crucial role in drug development, enabling drug developers to make dose recommendations for population subgroups without the need for expensive Phase 3 clinical trials. A central postulate in pharmacokinetics is that if two concentration-time profiles for a drug are identical across two different formulations or across two different occasions for the same formulation, the efficacy and safety should be equivalent (see Figure 2). This postulate forms the basis for generic drug development, formulation changes, the examination of drug interactions or food effects, and comparisons across subpopulations, such as patients with renal or hepatic impairment. Think about this: drug development would be prohibitively expensive if drug developers were required to do lengthy Phase 3 clinical trials in every population to show efficacy and safety for every subgroup or with every drug interaction.

Pharmacokinetic Equivalence Postulate

The fundamental assumption of bioequivalence is that if two drug products exhibit the same concentration-time profile, they will have essentially the same efficacy and side effects.



Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

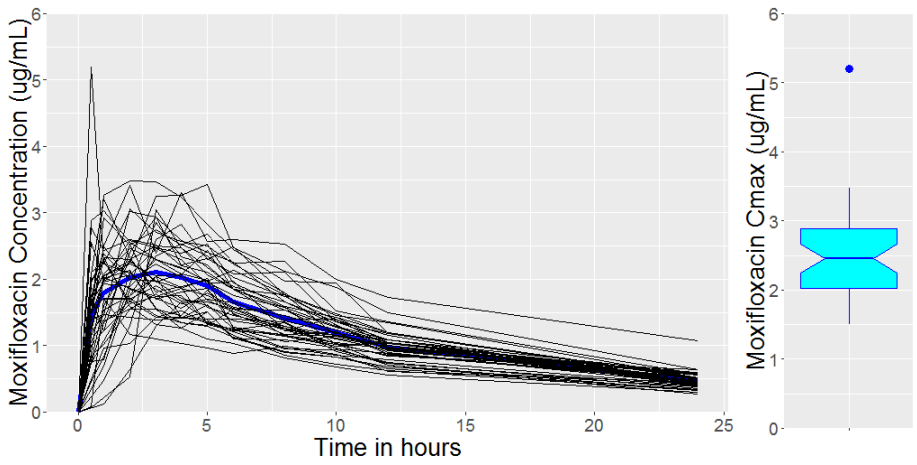


Figure 35. Example of variability in pharmacokinetics. Left plot: concentration-time profiles after single dosing in 43 individuals after receiving a single dose of 400 mg moxifloxacin. Notice the one individual whose maximal concentration exceeded 5 μg/mL, which is considerably higher than that of other subjects in the cohort. The solid blue line is the geometric mean concentration. Right plot: box and whisker plot of maximal moxifloxacin concentrations in the subjects in the left plot. The coefficient of variability in C_{max} was 28%, which is quite small for most drugs.

BE is assessed by comparing AUC, Cmax, and Tmax for each product or condition. Even for the same formulation of a drug under the same conditions, because of day-to-day differences, it would be impossible to demonstrate exact equality in AUC and Cmax. Hence, it is assumed that if the concentration-time profiles differ by ±20%, they are essentially equivalent. For example, Olling et al. [124] present the results of a BE study of three competitor formulations of carbamazepine (A, B, and C) and the brand-name drug (D). The results are shown below:

BE Tests for Carbamazepine Formulations as Reported by Olling et al [124]

	AUC			Cmax		
Test vs Reference	**A vs D**	**B vs. D**	**C vs. D**	**A vs. D**	**B vs. D**	**C vs. D**
Mean Ratio (T/R)	0.82	0.97	0.99	0.70	1.23	1.34
90% CI	0.73-0.92	0.86-1.09	0.88-1.09	0.60-0.83	1.04-1.46	1.13-1.58

Legend: CI, confidence interval. Reported is the geometric mean ratio, which is the ratio of the geometric mean of the test formulation to the geometric mean of the reference formulation after logarithmic transformation of the data. A value of 1 indicates the means are equal. A value < 1 indicates the test mean is smaller than the reference mean, and a value > 1 indicates the test mean is larger than the reference mean. The reference formulation is often the brand name drug, and the reference formulation is the generic drug.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Bioequivalence is declared if the 90% confidence interval for the geometric mean ratio for *both* AUC and Cmax is entirely contained within the interval of 0.8 to 1.25. Even though the confidence intervals for the mean AUC ratio for Formulations B and D ranged from 0.80 to 1.25, none of the formulations was bioequivalent to the brand-name product, as Cmax was not met for all three competitor products. Dizziness is a common side effect of carbamazepine. Formulations B and C had higher Cmax than the brand name drug. Subjects in this study who received Formulations B and C had a higher incidence of dizziness than the brand name drug, whereas Formulation A, which had a lower Cmax compared to the brand name drug, had a lower incidence of dizziness. Hence, in this study, differences in the formulations were directly translatable to carbamazepine's adverse event profile.

The pharmacokinetic equivalence postulate serves as the basis for dose modifications in specific subgroups and contraindications when a drug is coadministered with another. For example, if the exposure in patients with severe renal impairment was twice that of healthy patients, then a 50% dose reduction may be needed to ensure the patients with severe renal impairment have similar exposures as patients with normal renal function. It would be expected that the safety and efficacy of the two groups would be the same, despite one group receiving half the dose. This concept will be further discussed in the chapter on Special Populations.

Pharmacokinetic Variability

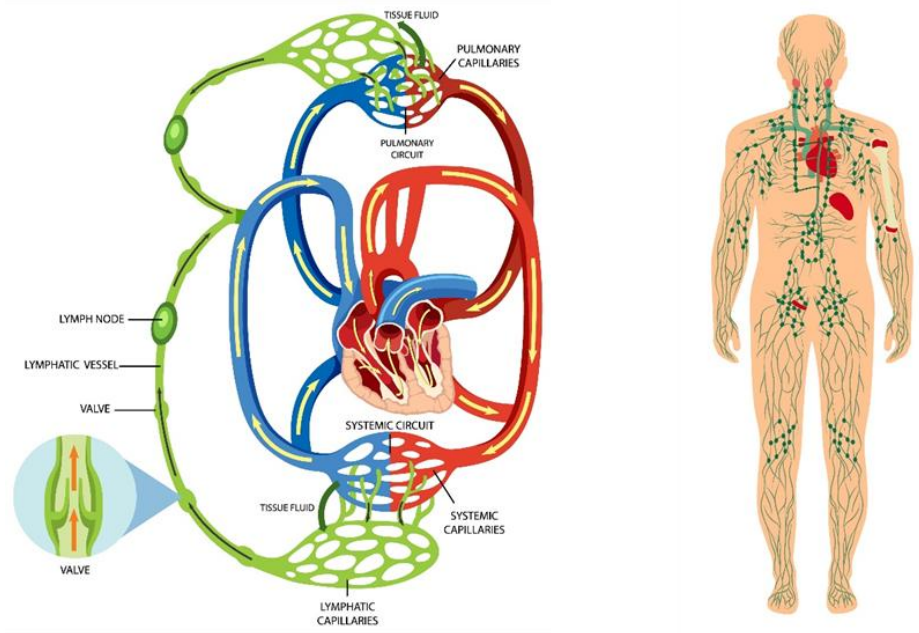
Population pharmacokinetics is the study of pharmacokinetics at the population level rather than the individual level. Its goal is to identify and characterize the sources of variability in a population and to identify patient-related factors, such as disease state, weight, or age, that might explain the variability. It does so through complex mathematical and statistical models.

So far, when pharmacokinetic parameters have been presented, they have been presented as a single value, e.g., a drug's half-life was 36 hours. It's important to remember that everyone is different. Hence, each individual will have different pharmacokinetics, and their concentration-time profiles will be different (Figure 35), and they will have different exposure than others. This variability is referred to as between-subject variability (BSV) and, for some drugs, can be quite substantial (>100%). The typical BSV for drug concentrations across patients is approximately 50%. The cause of BSV is myriad and often unknown. It may be due to genetics, intrinsic factors such as age or weight, extrinsic factors like smoking or diet, or other unknown factors. Depending on the exposure-response relationship (discussed in the next chapter), some patients may need higher or lower doses than others. For drugs with very high variability, it may be challenging to dose these drugs appropriately and may require therapeutic drug monitoring.

It should also be noted that within a subject, their pharmacokinetics may vary randomly from day to day or week to week. In other words, one day their C_{max} may be 100 ng/mL, and then the next time they take the drug, their C_{max} may be 150 ng/mL, and the next time it may be 84 ng/mL. This type of variability is referred to as within-subject variability (WSV) or inter-occasion variability (IOV). It is better to have a drug with a large BSV than a large WSV because a doctor may need to titrate a patient to the right concentration to achieve the desired drug effect. However, once they reach that dose, their concentrations remain stable and don't change over time. However, for a drug with a large WSV, its concentrations may change so much from day to day that it is hard for them to remain within the window of concentrations required to have a therapeutic effect. Typical WSV is around 20-50%. Causes of wide WSV may include diet, drug administration at different times of day, timing of drug administration relative to a meal, or the drug's poor formulation that doesn't disintegrate into solution consistently.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Diagram of the Lymphatic System



On the left is a schematic illustrating how the lymphatic system shadows the blood circulatory system, where all lymph fluid ultimately drains into the heart. On the right is an illustration of all the lymphatic veins in the body. In red, the important organs of the lymphatic system are highlighted: the thymus, spleen, marrow, lymph nodes, and tonsils. The function of the lymphatic system is to maintain fluid balance, aid in the absorption of dietary fats, and support the immune system by transporting lymph, filtering pathogens, and facilitating immune cell circulation. The lymphatic system is also important in the subcutaneous absorption of protein drugs.

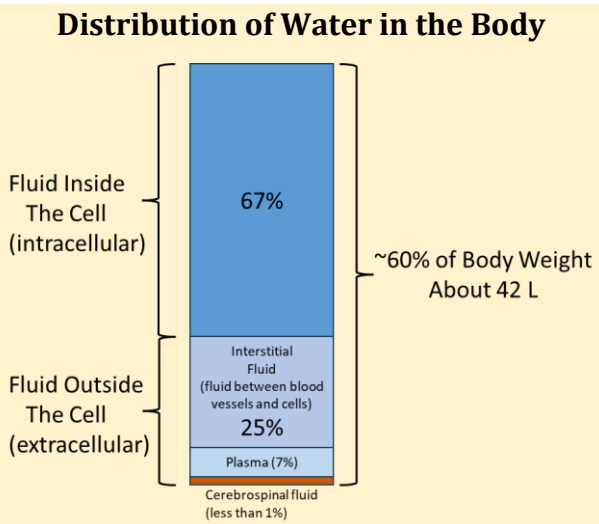
Pharmacokinetics of Biologics

Biologics are a heterogeneous class of drugs that range from hormones to monoclonal antibodies to vaccines. They differ from small molecules in terms of their synthesis, with chemical versus biological synthesis, and exhibit some interesting differences in their pharmacokinetics. However, having a general discussion about biologics is like trying to talk about motorized vehicles, from a Vespa to a Lamborghini. Nevertheless, a general discussion of the pharmacokinetics of biologics is warranted due to their unique properties compared to small molecules.

Biologics have molecular weights many times those of small molecules, and their physical size can be comparable to that of a skyscraper rather than a dollhouse. For example, the molecular weight of aspirin is 180 Daltons, whereas the molecular weight of the monoclonal antibody Humira is 148,000 Daltons, almost 1000-times larger. Due to their size and charge, which preclude absorption through the gut wall, and their denaturation at the low pH of the stomach, they cannot be administered orally. As such, biologics are predominantly administered by the intravenous, subcutaneous, or intramuscular routes. The intravenous route of administration provides complete bioavailability, whereas the other routes result in incomplete absorption. Biologics administered subcutaneously or intramuscularly have slow absorption, and those with molecular weights greater than 16 kDa are absorbed primarily via the lymphatic system.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Everyone is familiar with the circulatory system, which involves blood, but the lymphatic system is less well-known. The lymphatic system has been called our body's sewage system because it is a shadow circulatory system that parallels the blood circulatory system. The lymphatic system is considered part of the immune system because it protects us from infection by collecting fluid leaked from the circulatory system, producing certain cells and antibodies necessary for proper immune function. Compared with the circulatory system's blood flow, which has a flow rate of 5 L/min, the lymphatic system's flow rate is much slower, about 4 L/day. Due to their slow absorption rate, subcutaneously administered monoclonal antibodies take approximately a week to reach peak blood concentrations. Notably, they exhibit a bioavailability of 50% to 80% [125].



The figure illustrates the distribution of water in the body. Most water is found inside the cells. Monoclonal antibodies tend to remain mostly confined to the extracellular space. Antibody fragments and bispecific antibodies may have a larger distribution because of their smaller size.

For some biologics, particularly monoclonal antibodies, because of their large size, their distribution is limited to the body's interstitial and plasma fluid compartments. Their distribution into cells and the intracellular fluid is limited. As such, the volume of distribution of these drugs is typically small, ranging from 5 to 10 L. Furthermore, most biologics show no protein binding (consider that the molecular weight of a monoclonal antibody is 3-fold that of albumin). Still, some hormones, such as growth hormones, have binding proteins that regulate their levels in the body. Some new classes of biologics, such as monoclonal antibody fragments, differ from traditional biologics in that their molecular weight is significantly lower. As a result, they may have greater tissue distribution and larger volumes of distribution compared to traditional biologic drugs.

The metabolism of biologics is typically non-cytochrome P450-mediated and occurs through the same catabolic pathways as endogenous macromolecules. Proteins are metabolized by nonspecific proteases, and DNA-derived products are metabolized by nucleases. Depending on their size, biologics may or may not be renally eliminated. Most monoclonal antibodies fail to show any renal elimination; however, many hormones do, as their molecular weights are below the 60,000 Da cut-off in the glomerulus.

Although first described by Gerhard Levy in the 1990s to describe the kinetics of warfarin [126], many biologics have since been shown to have an unusual method of clearance relative to small molecules – a phenomenon known as receptor- or target-mediated drug disposition (TMDD). This capacity-limited system is another elimination mechanism to remove biologics from the body. Under the mechanism, the drug binds to its receptor, forming a drug-receptor complex. The body then clears this drug-receptor complex at a greater rate than it clears the drug itself. Because the number of receptors is finite, TMDD is a saturable, nonlinear system. Hence, the kinetics of drugs that undergo TMDD is complex. Drugs that undergo TMDD include efalizumab (Raptiva), omalizumab (Xolair), and abciximab (ReoPro).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Story of Epoetin Alfa and Antibody-Mediated Pure Red Cell Aplasia

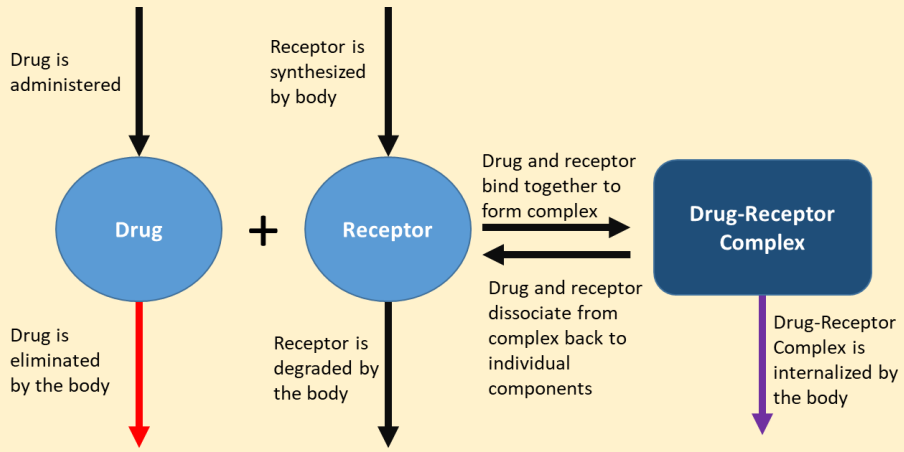
In 1999, Nicole Casadevall and colleagues in Paris identified in patients taking epoetin alfa (Epogen) for the treatment of chronic renal failure 13 cases of pure red cell aplasia (PRCA), a rare condition associated with anemia, low reticulocyte count, and low red blood cell precursors in the bone marrow. Serum from all patients blocked erythroid colony formation in bone marrow, an effect reversed by exogenous erythropoietin. In 12 of 13 patients, antibodies to erythropoietin were detected; however, these declined after erythropoietin administration was stopped. These antibodies blocked or neutralized the effect. Antibodies against a biologic drug are called 'anti-drug antibodies.' Anti-drug antibodies that negate the drug's effect are termed 'neutralizing antibodies.' In 2002, the Food and Drug Administration published a report of 82 cases of PRCA. Epogen, Procrit, and Eprex were all marketed products of epoetin alfa in the United States, with Epogen and Procrit sharing the same formulation. Eprex was different and was sold only outside the United States. Of the 82 cases of PRCA, 78 were in patients taking Eprex. In 2005, Johnson & Johnson, the manufacturer of Eprex, confirmed an additional 215 cases of antibody-positive PRCA, of which 189 were reported following Eprex administration.

An extensive investigation was made to explain the increased risk of developing neutralizing antibodies to Eprex with repeated administration. The exact cause has yet to be determined. Subcutaneous administration was identified as a risk factor, and in 2002, European authorities contraindicated its use in patients with chronic kidney disease. However, the major risk factor appeared to be Eprex's formulation. In the 1990s, Europeans were concerned about the risk of developing "mad cow" disease from products derived from animal sources. In 1998, albumin was removed from the Eprex formulation and substituted with polysorbate 80. Eprex was the only formulation containing human serum albumin; Procrit and Epogen contained other surfactants. It is believed that Polysorbate 80 forms micelles (lipid molecules that arrange themselves into spherical capsules in aqueous solutions), which could trap monomers or epimers of epoetin on the micelle surface, thereby increasing their antigenicity. Researchers also found that leachates from uncoated rubber vial stoppers increased antibody production, possibly by acting as an adjuvant to the immune response. Today, with changes in the route of administration and the use of coated rubber stoppers on all marketed products, the incidence of PCRA has decreased to very low levels [127].

As an example of TMDD, consider the pharmacokinetics of alemtuzumab, which was originally used for treating B-cell chronic lymphocytic leukemia (Campath), and today is used to treat multiple sclerosis. In leukemia patients, alemtuzumab was escalated up to 30 mg three times weekly and was continued for 12 weeks. After the first 30 mg dose, alemtuzumab's half-life was 11 hours, but after 12 weeks of therapy, the half-life increased to 6 days. Alemtuzumab binds to CD52 receptors on B- and T-cells and, upon binding, causes cell death via antibody-mediated cell lysis. As the cell dies, the receptors available for alemtuzumab binding are lost.

Hence, when a patient starts alemtuzumab, high receptor levels are observed due to the patient's leukemia, and faster clearance occurs due to the formation of the drug-CD52 complex and cell lysis. As the leukemia is cleared, there is a progressive loss of available CD52 Receptors, and alemtuzumab clearance decreases, leading to an increase in its half-life. Similar results were seen with rituximab (Rituxan), a monoclonal antibody used to treat some cancers. In leukemia patients, the half-life after the first dose is ~3 days, but after 4 weeks of treatment, it increases to ~9 days [128].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Target-mediated drug disposition (TMDD) was originally called receptor-mediated drug disposition by Gerhard Levy in 1994. Later, it was changed to TMDD to reflect that drugs can also interact with non-receptor targets. With TMDD, another clearance mechanism occurs for the drug. As the drug binds to its target receptor, a drug-receptor complex forms, which the cell internalizes, removing the drug from the systemic circulation. Hence, the drug is cleared by linear or nonlinear elimination (red arrow) and internalization (purple arrow).

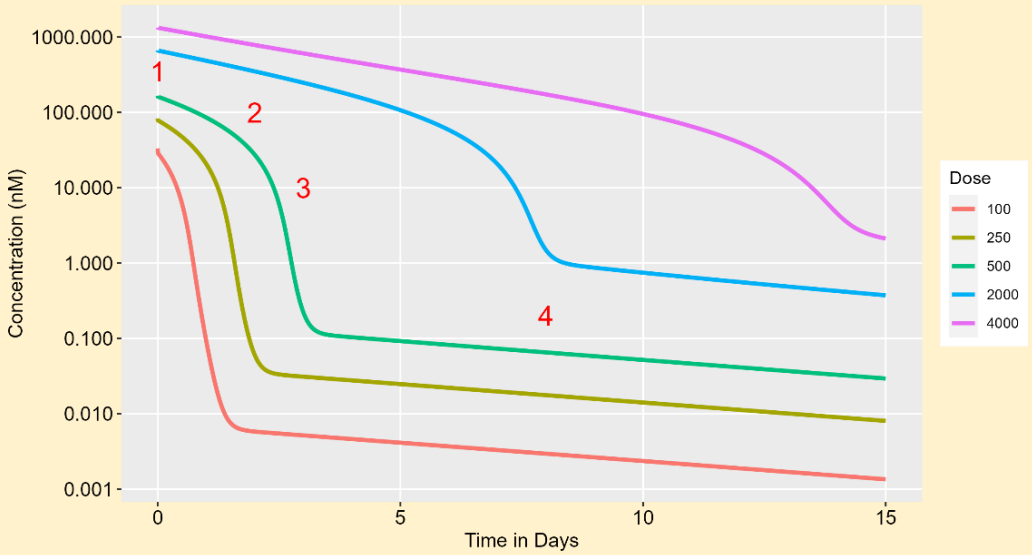
Another important aspect of biologics is their immunogenicity. Early biologics, like insulin and the first monoclonal antibodies, were derived from animals. When administered to humans, the body treats foreign proteins as foreign and mounts an immune response against them. These antibody-drug complexes could affect a drug's pharmacokinetics, as well as its efficacy and safety. Even as biologics were manufactured using recombinant DNA methods and became increasingly "human" in nature, the body often continued to mount an immune response against them. Today, antibody formation against biologics is a major concern for drug developers and regulatory agencies. Factors that influence the degree of immunogenicity of a biologic include the route of administration (subcutaneous is more immunogenic than intravenous), formulation (see the side box on The Story of Epoetin Alfa), dose (higher doses lead to greater immune response), length of treatment, impurities in the formulation, sequence variation of the biologic ("how human is it?"), degree of glycosylation of the protein, patient-related factors, and factors that are still unknown [129].

Pharmacokinetics of Cell Therapies

The pharmacokinetics of cell therapies differ from what has been presented and warrant examination due to their unique characteristics. Recall that cell therapies are either stem-cell-based or non-stem-cell-based, and if they are the latter, they are either immune cells or non-immune cells. Further, they are either autologous (from the same patient) or allogenic (from another patient). They can be unmodified or modified. Since CAR-T cells are the most prominent example of cell therapy, we will focus on them for this discussion. The first approved CAR-T therapy was tisagenlecleucel (Kymriah), an autologous CAR-T cell therapy targeting CD19+ B cells, approved in 2017 for certain types of leukemia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Pharmacokinetic Profile of a Monoclonal Antibody Showing Target-Mediated Drug Disposition



Drugs whose pharmacokinetics are influenced by their binding to a receptor or target show "target-mediated drug disposition", or TMDD for short. Monoclonal antibodies are a classic example. Their concentration-time profiles are like a crazy children's slide as the drug binds to and releases from the target.

Up to 4 phases can be seen after rapid intravenous administration:

1. a steep, initial descent due to rapid binding to the receptor;
2. a linear elimination phase, where the target is saturated with drug and no more binding occurs, and elimination is via the usual clearance mechanisms;
3. a mixed-order transition phase where the target is partially saturated, and the drug starts to bind to the target again, and
4. the terminal elimination phase, where the drug's elimination occurs only by internalizing the target-receptor complex, thereby shifting the equilibrium back to the free, unbound drug form and allowing the drug to bind the target again.

Not all drugs may exhibit all phases; see, for example, the purple line in the figure. Whether this occurs depends on how tightly the drug binds to its target and how fast the internalization rate is relative to the usual clearance mechanisms [130].

Further, TMDD can occur not only with monoclonal antibodies but also with small molecules. Some ACE inhibitors and nicotine have been proposed to display TMDD.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

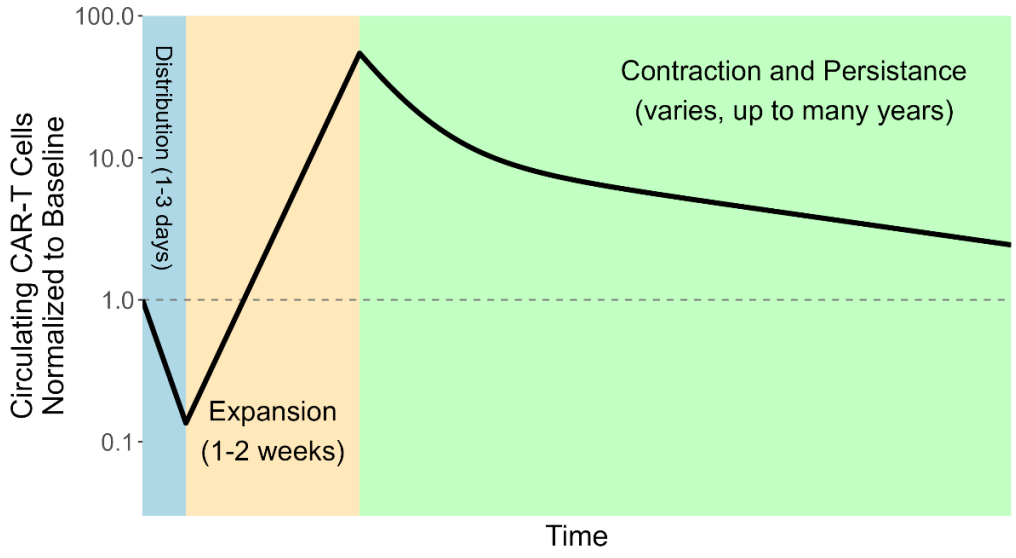


Figure 36. CAR-T cell pharmacokinetics.

After CAR-T cell administration, there are three distinct phases in cellular pharmacokinetics (Figure 36)[131]:

- The distribution phase, where the drug distributes to the tissues and cell counts decline; this takes 1-3 days.
- The expansion phase lasts about 1-2 weeks. During this period, cells undergo clonal division and expansion, resulting in a 100-fold increase in their number. The maximum number of observed cells is called Cmax, similar to the concept in small molecule pharmacokinetics.
- The contraction and persistence phase. Cell counts stop increasing and begin to decline due to programmed cell death, a process called the contraction phase. Thereafter, cell counts remain relatively stable for years, a period known as the memory or persistence phase.

Efforts to understand whether any patient characteristics can affect Cmax values have largely been unsuccessful to date [132]. However, Cmax values can vary by disease type, and higher Cmax values are associated with a better clinical response – the higher the Cmax, the more likely the patient is to respond to the drug. There is also some evidence that greater pretreatment tumor burden can result in larger clonal expansion. To prevent cytokine release syndrome, patients are often pretreated with corticosteroids. These drugs do not appear to affect Cmax.

Flat Dosing or Weight-Based Dosing?

Some drugs are administered as a flat dose - everyone gets 100 mg for example. Other drugs are administered weight-based, such as 10 mg/kg or 100 mg/m2, where the administered dose depends on the patient's weight or another weight-based metric, such as BSA or BMI. A modification of weight-based dosing is dose banding, in which doses are rounded to a fixed value if the calculated dose falls within a specified range. Dose-banding, which is more common in the EU than in the US, is proposed to make dosing easier than actual weight-based dosing and can reduce unused drug waste. For example, suppose the dose was 1000 mg/$m^{2,}$ and there was a 500 mg vial strength. A dose band may be as follows:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Proposed Dose Bands		
BSA (m^2)	**Calculated dose (mg)**	**Actual Dose (mg)**
1.4 to 1.8	1400 to 1800	1500
1.8 to 2.3	1801 to 2301	2000
2.3 to 2.8	2301 to 2800	2500
2.8+	2801 or greater	3000

Weight-based dosing has its origins in oncology, where dosing of early chemotherapeutic agents was based on the patient's weight or BSA. Some drugs with high renal clearance, such as carboplatin, are dosed based on the patient's glomerular filtration rate (GFR), which is estimated from the patient's weight. Based on the theory that larger patients have larger volumes of distribution and larger metabolic capacity, these patients should require larger doses to compensate and reach similar concentrations as smaller patients [133]. Further, early animal studies of cancer drugs used weight-based dosing to allow for allometric scaling across species; e.g., it was assumed that 50 mg/m2 in rats was equivalent to 50 mg/m^2 in humans. Hence, for decades, weight-based dosing has been the *de facto* standard for adults. As pharmacokinetics became more important in the 1980s and 1990s, it became apparent that the real goal should be to find a dosing regimen that minimizes between-subject variability across a wide range of body weights. When viewed in that light, many began to question whether weight-based dosing was really doing what people thought it was, and whether it was appropriate or even needed.

With biologic drugs, the situation is even more confusing: within the same class, some may be weight-based and others flat-dosed, and even within weight-based drugs, there is no agreement on the weight metric. SyBing and colleagues from Pfizer [134] examined 143 biologic drugs identified through FDA databases and extracted whether they were flat-dosed or weight-based. Of the 77 monoclonal antibodies found, 32% were dosed based on weight, whereas 61% were flat-dosed. Early development of monoclonal antibodies was largely weight-based, but today, that is not necessarily the case. For new drug classes, like bispecific antibodies, which have only been around a decade or so, 56% were flat-dosed, whereas 44% had weight-based dosing. For antibody drug conjugates (ADCs), it's generally agreed that they should be weight-based in their dosing due to their toxicity and narrow therapeutic window, yet among the 12 ADCs identified, 75% were dosed mg/kg, whereas 17% were dosed mg/m^2. Why the difference? Often, there is no good scientific rationale and the choice is made based on preconceived notions and historical reasons, i.e., others did it this way, we should do it this way too.

For many drugs, there is a positive relationship between clearance and weight-based metrics. As these metrics increase, so does clearance. This relationship can be described using a power function:

$$\text{Clearance} = CL_T \left(\frac{\text{body size}}{\text{Typical body size}} \right)^{\alpha} . \tag{19}$$

Body size metrics may include weight, BSA, ideal body weight, fat-free mass, and many other metrics. Usually, weight is modeled, and the typical body size is set to 70 kg; CL_T is then the clearance for a person having the typical body size. SyBing and colleagues [134] propose that if α is < 0.7 for a drug, then fixed weight dosing is needed, and if α is > 0.7, then weight-based dosing is needed. Essentially, when there is a modest relationship between clearance and weight, such as when is $\alpha < 0.7$, the variability in observed drug concentrations across patients obscures any such weight relationship. This is not to say there is no relationship; it's just that the noise in the data is such that the relationship is not apparent. It is only when the relationship is strong, such as when is $\alpha > 0.7$, that the weight relationship can be seen through the noise of between-subject variability.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

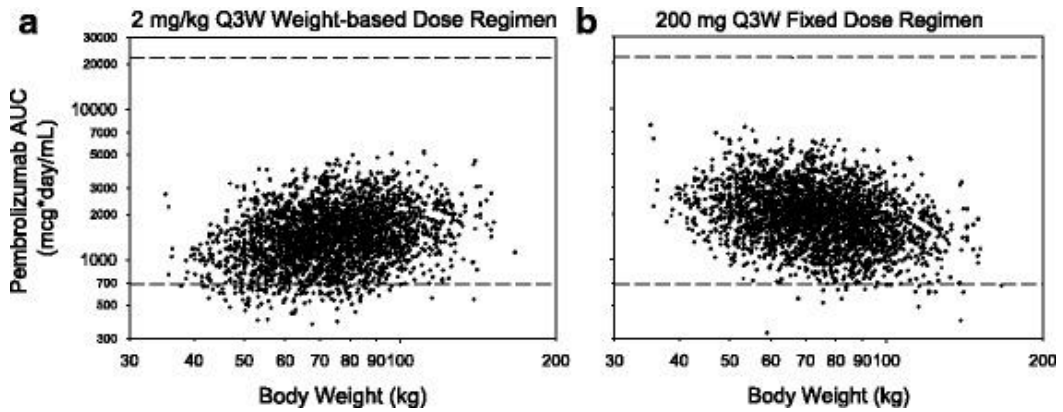


Figure 37. Predicted variation in pembrolizumab AUC exposure by body weight for weight-based (**a**) and fixed-dose (**b**) regimens (2800 replicate simulations): *Horizontal dashed lines* represent the range of exposures (5th percentile of 2 mg/kg Q3W and 95th percentile of 10 mg/kg Q2W) from dose regimens demonstrated to have comparable efficacy and tolerability in melanoma and NSCLC trials. Regulators concluded that there was no difference in AUC values overall, allowing Merck to change its label from weight-based to fixed dosing. Figure reprinted from [135].

Today, there is an increased push towards fixed dosing in adults, as studies show no real advantage to weight-based dosing. For example, Keytruda (pembrolizumab) is one of the top-selling cancer drugs, and it is flat-dosed. For adults, the dose is 200 mg every 3 weeks or 400 mg every 6 weeks. But it wasn't always this way. When it was approved in 2015, the adult dose was weight-based: 2 mg/kg as a 30 minute infusion. Pembrolizumab has a power exponent α of 0.58, which, according to SyBing et al., should use flat-based dosing. Pharmacokinetic simulations, and confirmation using observed data collected in a limited number of patients using a flat dose, show that there is no difference in drug concentrations when the drug is dosed 2 mg/kg every 3 weeks or 200 mg every 3 weeks (Figure 37) [135]. In 2016, the label was changed from weight-based to fixed dose, where it remains today.

For children, however, weight-based dosing remains the standard. In fact, drugs that are flat-dosed may switch to weight-based dosing in children. For example, returning back to pembrolizumab, the adult dose is a fixed dose, but the pediatric dose is 2 mg/kg, which is weight-based.

Summary

Pharmacokinetics is the foundation for how a drug works. A drug must be absorbed and reach the site of action at a sufficient concentration to exert an effect. Once at the site of action, the drug must bind to a receptor and have an effect. These three steps form the basis of the "three pillars." The next chapter will discuss pharmacodynamics, examining what happens when a drug binds to its receptor to elicit an effect.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 5

Principles of Pharmacodynamics

The previous chapter discussed pharmacokinetics, or what the body does to the drug. However, considering pharmacokinetics alone in explaining Drugs Explained is only half the story. The other half of the story is a drug's pharmacodynamics. It's like talking about Batman without Robin, Laurel without Hardy, or peanut butter without jelly. It's simply incomplete. Pharmacodynamics is the effect a drug has on the body. Pharmacodynamics is the yang to the pharmacokinetics' yin. While the body is trying to remove the drug from the body, at the same time, that drug is doing something to the body. This chapter introduces pharmacodynamic principles, focusing mostly on small-molecule drugs, and explains how pharmacokinetics and pharmacodynamics interact with each other.

Drug Targets

How many drug targets there are in the body is debatable, with estimates ranging from the hundreds to the thousands [136, 137].

A drug target is a broad concept that encompasses the idea that a drug binds to a specific biomolecule, altering or changing its function. Early in the history of pharmacology, the primary drug targets were primarily proteins, such as receptors and enzymes. Today, there are many drug targets, which are still mostly proteins, but also include DNA and transcription factors. Additionally, some drugs don't have a specific target, such as cell and gene therapies. In a recent analysis of first-in-class drugs approved by the FDA between 2019 and 2022, 42% of approved drugs targeted classic protein mechanisms, 22% targeted soluble proteins or peptides, and the remainder employed other mechanisms [138]. Between 2011 and 2022, over 60% of approved drugs utilized classic protein mechanisms. Hence, over the past 10 years, there appears to be a shift away from traditional targets to more nontraditional ones.

A key concept in the idea of a target is "drugability," or whether the target is "druggable." While a target may be identified, creating a drug that can reach or interact with it may not be possible. Some targets are undruggable. Examples of currently undruggable targets include tyrosine and serine phosphatases, as well as transcription factors such as p53 or Myc. It is estimated that 85% of all potential drug targets are undruggable. However, today's undruggable targets are opportunities and may become tomorrow's new drugs. For a long time, KRAS, a protein involved in cell growth and one of the most commonly mutated cancer genes, was considered undruggable. But in 2021, after a lot of research, Lumakras (sotorasib) was approved for the treatment of non-small cell lung cancer and is a KRAS inhibitor. There are many other opportunities for making undruggable targets druggable.

A good drug binds to its target with high specificity and high affinity, meaning it binds to only the target and exhibits a strong attraction to it. Most drugs, however, don't bind to a single, specific target. Even the most specific drugs for a target bind to other biomolecules to some degree (Figure 38), and when these other targets are modulated and show pharmacologic effects, these are referred to as "off-target" effects. One estimate suggests that more than half of all small molecules bind to

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026.

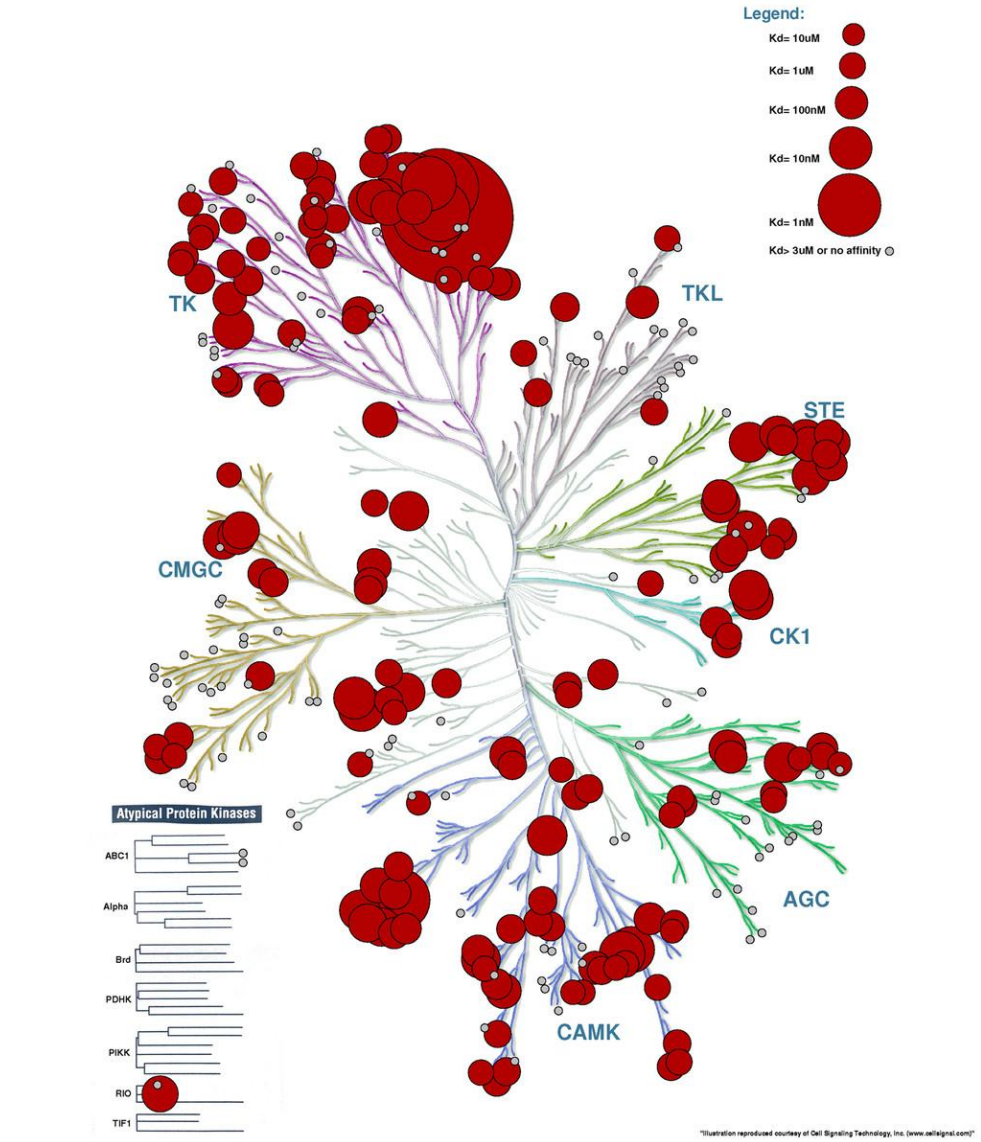


Figure 38. The human kinome displays sunitinib's binding affinity (Sutent). Each branch and arm in the figure represents a unique human kinase. Every branch is a potential target for a drug. The red circles indicate sunitinib's antagonist affinity for these different kinases. Notice that sunitinib inhibits more than 80 kinases. The larger the red circle, the greater sunitinib's affinity for blocking these kinases. Figure is reprinted from [139].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

more than five proteins [140], but this figure has been criticized as possibly too low because of the limited number of protein target assays that most companies test during drug development [141]. Typically, off-target effects are undesirable and can lead to adverse events.

For example, terfenadine (Seldane) was an antihistamine, i.e., it was an antagonist at histamine receptors, that was used therapeutically to treat seasonal allergies, but it also blocked potassium ion channels in the heart, which in some patients led to fatal arrhythmias and was the reason for its removal from the market in 1997. But not all off-target effects are bad. The beneficial effects of many drugs result from polypharmacology, or the binding of a drug to multiple targets that collectively produce the desired benefit. Aspirin is one such drug. It is believed that aspirin binds to more than 20 proteins and modulates several different biochemical pathways [142]. Polypharmacology, through the administration of different drugs that affect different targets, called combination therapy, is also a mainstay of modern cancer treatment.

Distribution of Drug Targets of First-in-Class Drugs Approved by the FDA Between 2019–2022

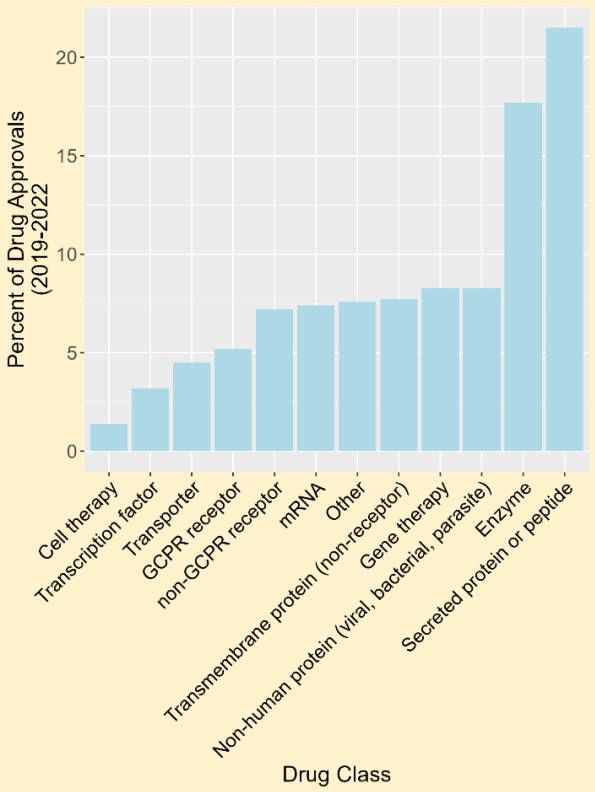


Figure redrawn from [138].

Receptor Theory

A receptor is a specific molecule in the body associated with some physiological process that, when interacted with by a ligand, drug, or stimulus, can bring about a cellular response. These include many cell membrane proteins, such as ion channels, metabotropic receptors (also known as G-protein-coupled receptors or simply G-proteins), cell membrane enzymes like the tyrosine kinase EGFR, transporters, soluble enzymes, and intracellular macromolecules like DNA or RNA. There are also other non-ligand-associated receptors like mechanoreceptors on the skin that detect touch, pressure, and vibration, photoreceptors in the retina that

Of Agonists and Antagonists

- Agonists are chemicals that bind to receptors and produce a biological effect. Endogenous agonists are the natural ligands that bind to a receptor.
- An antagonist blocks the effect of an agonist. Antagonists have no effect alone.
- Partial agonists are agonists that do not produce a full effect when the receptor is saturated (for example, buprenorphine used to treat opiate dependence).
- Inverse agonists are chemicals that bind to the receptor and produce the opposite effect of an agonist (for example, rimonabant).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

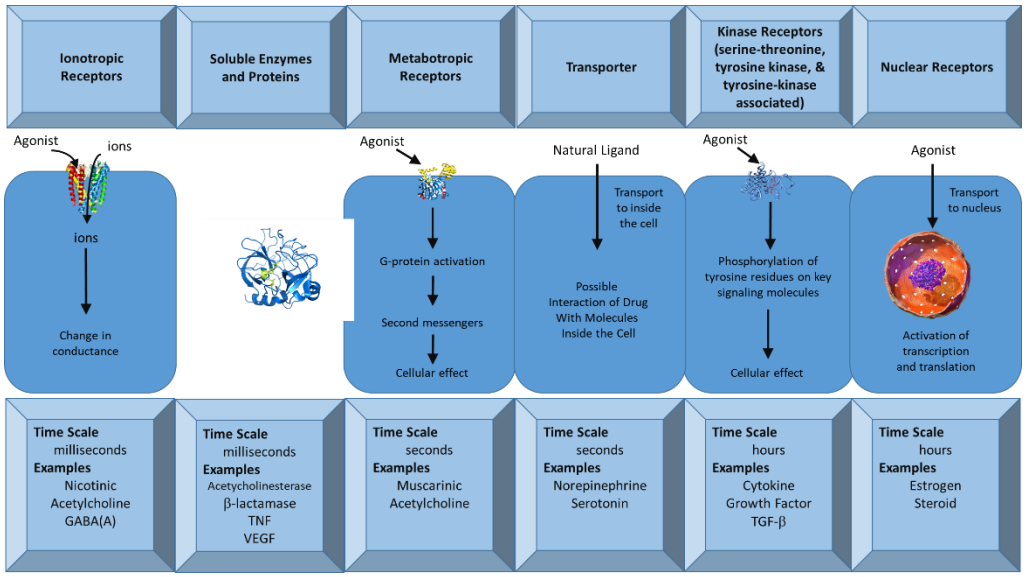


Figure 39. Major receptor superfamilies.

detect light, chemoreceptors in the nose and tongue that detect chemicals and mediate taste and smell, nociceptors in the skin and organs that detect pain, and thermoreceptors that detect temperature changes.

The receptor concept began in the early 1900s with the work of John Newport Langley (1852-1925), who proposed the 'receptive substance' concept based on his salivary secretion experiments in the dog, and by Paul Ehrlich (1854-1915), who flipped the receptor idea on its head and proposed the idea of a "magic bullet," in which it was possible to kill microbes specifically without harming the body using specific drugs for their target [144]. Both ideas are complementary and form the basis of pharmacology, which holds that drugs act specifically on receptors to elicit effects.

For almost fifty years, researchers expanded on and debated these ideas. They remained largely theoretical until 1960, when Peter Waser published autoradiographic pictures of curare binding to muscarinic receptors at the endplate of diaphragm muscle cells. These photos were exposed for months to allow the film to detect low levels of radioactivity. When completed, the experiments proved the existence of the receptors, which could be abolished by denervation of the muscle.

Autoradiographic image of curare binding to the diaphragm muscle. The dark spots are radiolabeled curare, which binds to muscarinic receptors. Photo reprinted with permission from [143].

As shown in Figure 39, there are five major cell membrane receptor superfamilies: ion channels, G-proteins, kinase receptors, transporters, and nuclear receptors [145]. These receptors have very different time scales upon activation. Ion channels work very fast, on the order of milliseconds, whereas nuclear receptors take hours to weeks to activate. They are also located in different regions of the body and cells. Ion channels, transporters, G-

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

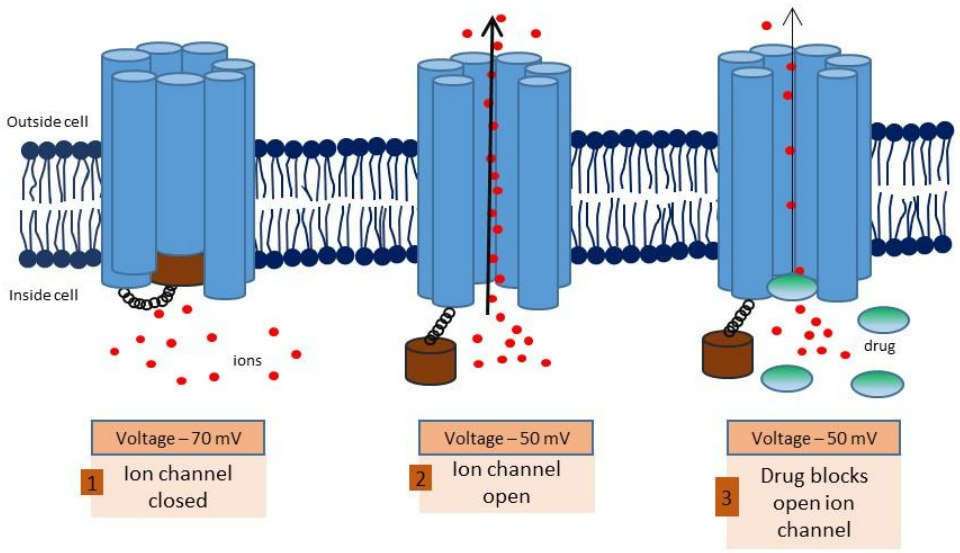


Figure 40. Schematic of a voltage-gated ion channel. Voltage-gated ion channels open and close in response to the electrical difference between the inside and outside of the cell. In the right-most two channels (#2 and #3), a subunit has been removed to illustrate the flow of ions through the channel.

proteins, and kinase receptors are on the cell membrane. Other kinases and enzymes are located in the cytoplasm, while some enzymes are found in the blood, and nuclear receptors are situated in the nucleus. They also have different effects.

Ion Channels: Ion channels and transmembrane proteins present in cells and some organelles are fast-acting and can be categorized into two types: voltage-gated (Figure 40) and ligand-gated (Figure 41). Their function is to control the cell's resting potential (the electrical charge difference between inside and outside the cell), modify action potentials, and regulate electrical signals within the cell by controlling the flow of ions through channels. A million ions or more can flow through an ion channel in a second, and they are in rapid equilibrium with plasma drug concentrations. Voltage-gated ion channels open or close in response to a change in the electrochemical voltage difference between the inside and outside of the cell. Examples of voltage-gated ion channels include sodium, potassium, chloride, and calcium channels. Ligand-gated ion channels open or close when a ligand, usually a neurotransmitter, binds to the channel. Examples of these receptors include GABA, glycine, ionotropic glutamate (AMPA and NMDA), serotonin (5HT-3), and nicotinic receptors.

Consider This

Consider a typical small molecule, having a plasma concentration of 1 ng/mL, which is equal to 1 part per billion; there are more than a thousand billion million molecules of the drug in 1 mL of plasma. Even for a drug 99.9% bound to plasma proteins, more than a billion million molecules of drug are available to bind to a receptor [146].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

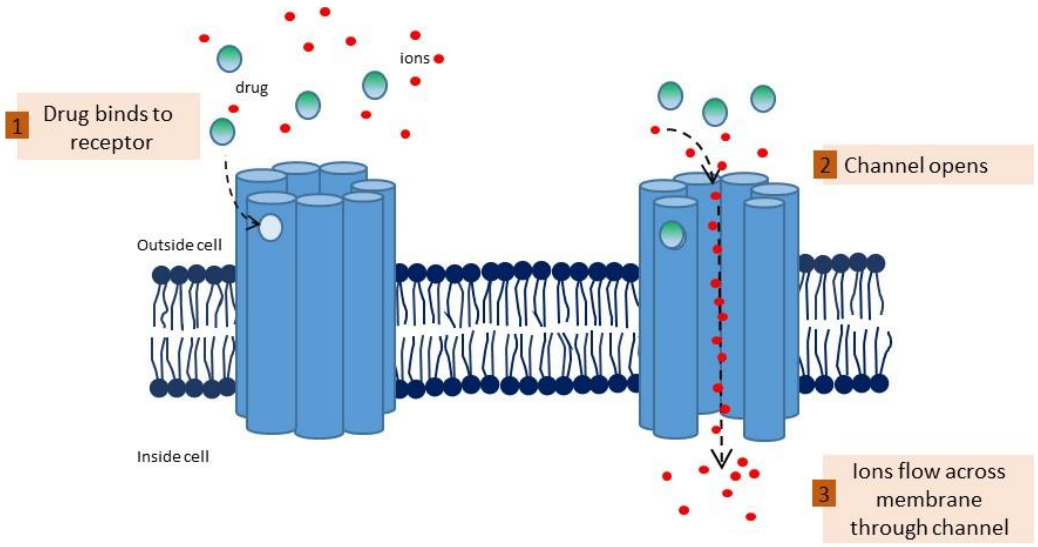
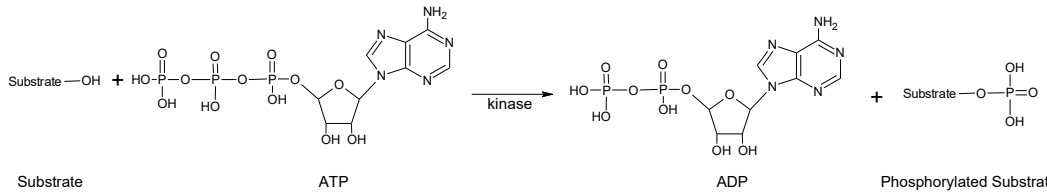


Figure 41. Schematic of a ligand-gated ion channel. Ligand-gated ion channels open or close in response to binding a ligand, often a neurotransmitter, to the receptor. In the right-most channel, a subunit is removed to illustrate the flow of ions through the channel.

Metabotropic or G-protein Receptors: Metabotropic receptors, also known as G-proteins, do not directly activate an ion channel but rather do so through intermediates called second messengers. These receptors are in a slow equilibrium with plasma concentrations because there is often a delay between peak drug concentrations and the drug's peak effect. The binding protein is often sent to a G-protein coupled receptor (GCPR), which is a seven-transmembrane-spanning receptor on the cell surface coupled to a G-protein that generates second messengers (see Figure 42). GCPRs are the largest family of human membrane proteins and are also among the most significant drug targets. G-proteins themselves have three subunits (Gα, Gβ, and Gγ) that generate hundreds or even thousands of molecules of second messengers, such as cyclic AMP (cAMP), which, when stimulated, activate other proteins, including ion channels. These receptors include metabotropic glutamate, muscarinic, and most serotonin receptors.

Kinase Receptors: Kinase receptors, which are located both on the cell surface and inside the cell, are enzymes that catalyze phosphorylation reactions, which is the transfer of phosphate groups from high-energy phosphate donors, like adenosine triphosphate (ATP), to form a phosphorylated substrate; the process of phosphorylation activates or "turns-on" the protein. This process is shown below, where ATP donates a phosphate group to the substrate, leaving behind adenosine diphosphate (ADP) and forming the phosphorylated substrate:



Through phosphorylation, kinases transmit signals from the outside to the inside of the cell and regulate complex cellular processes (Figure 43). There are many types of kinases, which are classified by the substrates they act upon. They are also often named by their specific substrates. Protein kinases are the most common drug targets and act by phosphorylating tyrosine, serine, threonine, or histidine amino acid residues. Although tyrosine and serine kinases are the most common protein

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

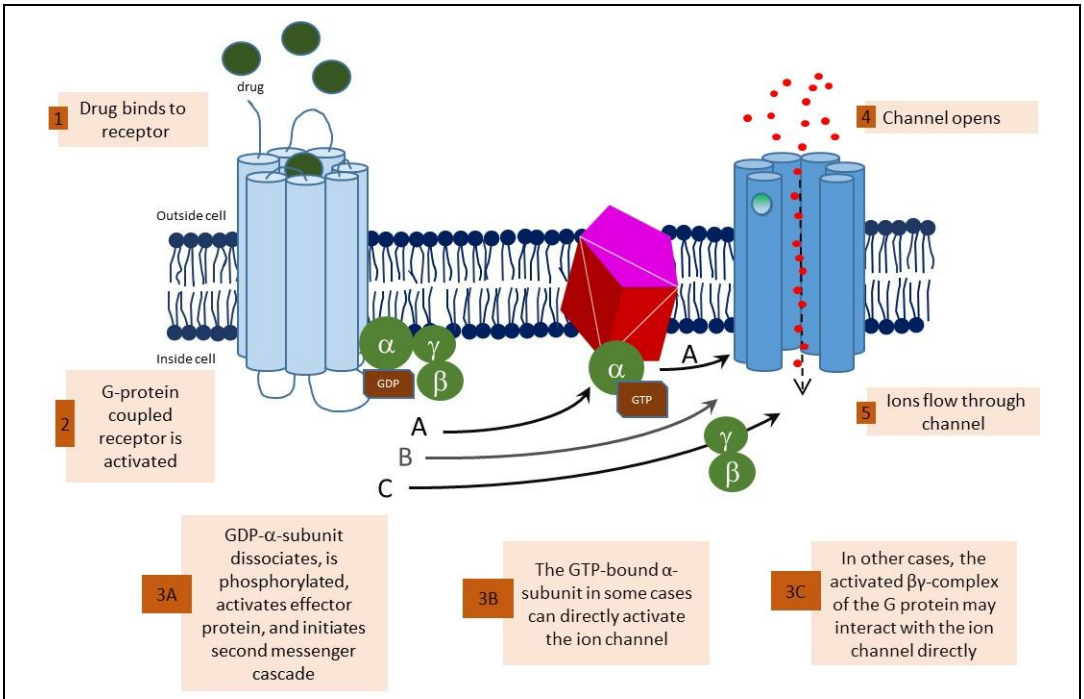


Figure 42. Schematic of a metabotropic G-protein coupled receptor (GCPR) that interacts with an ion channel. Metabotropic receptors act indirectly via ion channels by activating second messengers, such as cAMP. In the ion channel, a subunit is removed to illustrate the flow of ions.

kinases, there are others, like lipid and carbohydrate kinases. To date, more than 500 mammalian kinases have been identified. Because kinases act as intermediaries in cellular processes, they are slower to activate than ion channels, and their effects can take hours to years to develop (Figure 44).

Transporters: Transporters, unlike ion channels or G-proteins, are proteins that span the cell membrane and carry ions, peptides, small molecules, and even large molecules across the cell membrane, either from outside the cell to inside the cell or vice versa [147]. Whereas ion channels sit in the membrane and allow the flow of ions across the membrane, transporters bind the ligand and ferry it across the membrane one by one. Transporters are much slower than ion channels and can transport only a limited amount from one side to the other. Examples of transporters include P-glycoprotein, which transports bile and bile acids, and organic cation transporter-1 (OCT-1), which transports many neurotransmitters. Drugs can act as transporters' antagonists, blocking the transport of their natural ligands across the cell membrane.

Enzymes: Soluble enzymes are not bound to cell membranes [143]; some are in the blood, others are inside the cell. An example of a soluble protein is acetylcholinesterase, which is found at the neuromuscular junction, the space between muscle and nerve, that acts to break down acetylcholine, the major neurotransmitter involved with muscle contraction. Drugs can act as antagonists at soluble targets to decrease their activity or as replacement therapy to increase their activity. An example is the administration of recombinantly-derived imiglucerase (Cerezyme), an enzyme found inside the lysosome of cells, to patients with Gaucher disease, a rare disease in which patients do not make imiglucerase in their body and, as such, accumulate glucosylceramide, the substrate for imiglucerase, in various tissues, leading to anemia, enlarged liver, and bone problems. Technically, these are not receptors because they serve as replacement therapy; however, mentioning them in this section is useful. Soluble enzymes and transporters are unique because the concept of an agonist does not apply to these targets, as no drugs have been shown to act as an agonist at these targets. Drugs act only as enzyme and transporter antagonists.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

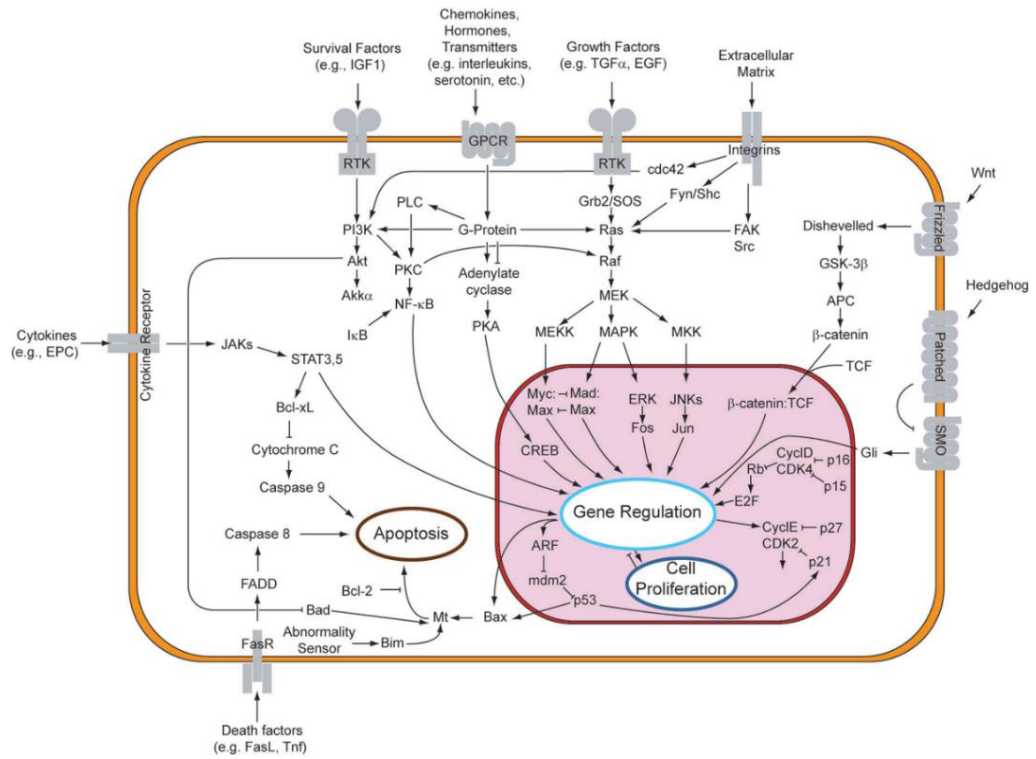


Figure 43. Overview of signal transduction pathways in the cell. Many of the proteins involved in these pathways are kinases.

Figure reprinted from Kinases, Wikipedia.

Nuclear Receptors: The last receptor superfamily is nuclear receptors. These are found in the cell's nucleus and regulate gene expression, thereby controlling cell development, metabolism, and homeostasis. Ligands for nuclear receptors are steroids, some vitamins, and hormones. Some nuclear receptors, known as orphan receptors, lack known endogenous ligands. What differentiates nuclear receptors from other receptor types is that their activation controls the expression of DNA in the cell.

Receptor Distribution: Receptors are not uniformly distributed throughout the body; they are localized in specific tissues and organs. Raymond Alquist showed in 1948 that there were two different types of adrenergic receptors (cell surface receptors that bind epinephrine and norepinephrine and play a key role in the central nervous system), which he named 'alpha' and 'beta.' Today, it is known, with the aid of very specific agonists and antagonists and the advent of molecular biology techniques, that most receptors exist as distinct subtypes. The same endogenous ligand can activate each subtype, but each subtype has distinct specificities for agonists and antagonists due to differences in amino acid sequences, and consequently, each subtype has distinct pharmacologic effects when modulated.

For example, dopamine receptors are a class of G protein-coupled receptors (GPCRs) found in the central nervous system and have dopamine as their natural ligand. In 1979, it was demonstrated that there are two dopamine subtypes, D1 and D2, which can be distinguished biologically, pharmacologically, physiologically, and by their location. Pharmacologically, D1 receptors bind the antagonist SCH-23390, whereas D2 receptors bind the antagonists spiperone and haloperidol. Physiologically, when activated, they exert their effects via distinct G-protein complexes: D1 activates cAMP signaling, whereas D2 inhibits cAMP production. And while their anatomical distributions overlap in the brain, the ratios of these subunits in different parts of the brain differ by location.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

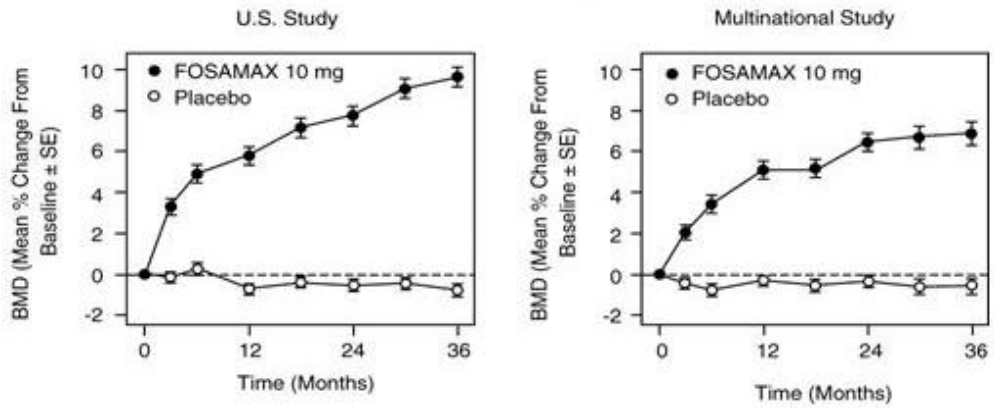


Figure 44. Sometimes, the effects of drugs take years to realize fully. Fosamax (alendronate sodium) is used to treat and prevent osteoporosis in postmenopausal women. After being taken into osteoclasts during bone resorption, Fosamax (alendronate) 10 mg/day inhibits the kinase enzyme farnesyl pyrophosphate synthetase. This figure shows that lumbar spine bone mineral density continues to increase in postmenopausal women years after they start taking the drug.

Figure reprinted from https://www.drugs.com/pro/fosamax.html.

For more than ten years, the two-subunit dopamine classification system remained in place until the emergence of molecular biology. In 1987, receptor cloning confirmed that the D2 receptor was a G-protein. Following the success of this method, others adopted a similar approach. The D1 receptor was cloned a few years later, and shortly after that, three new dopamine receptors were cloned: D3, D4, and D5. This is where we are today: dopamine receptors have five subtypes. These subtypes share similar, but not identical, amino acid sequences. The D1 and D5 subtypes share 79% amino acid homology but only 40% homology to the D2, D3, and D4 subtypes. However, the D2, D3, and D4 share homology levels ranging from 51% to 75%. Whereas the D1 and D5 stimulate cAMP production, no second messenger system has been identified for D2, D3, and D4 receptors. Receptor subtypes may have important pharmacological consequences because it may be possible to tailor drugs to a particular subtype, thereby decreasing the risk of adverse events.

Properties of Receptors

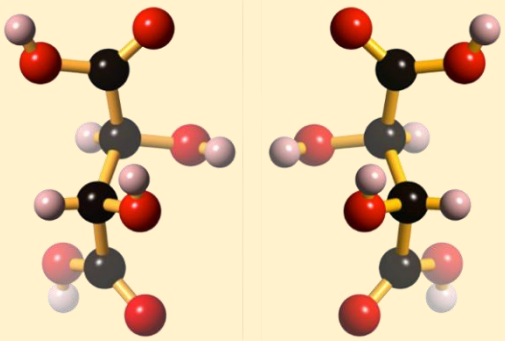
- **They are specific in their binding to agonists;**
- **They are saturable and finite in number;**
- **They have a high affinity for their natural ligands and**
- **Once a ligand binds to a receptor, the receptor does something or is prevented from doing something.**

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

How Important is Chirality?

Let's go back to 1848 in France, when a 26-year-old Louis Pasteur was working on his PhD in Chemistry and attempting to crystallize different compounds. A few years earlier, it was found that the sediment of wine casks contained high amounts of paratartaric acid. Although paratartaric acid behaved differently chemically from tartaric acid, it was found to have the same chemical structure as tartaric acid. This was paradoxical. How could compounds having the same chemical structure behave differently chemically? Pasteur sought to resolve this paradox. Shining a polarized light through an organic solution of pure acids, he found that tartaric acid rotated light to the right, but that paratartaric acid did not. He was puzzled. How was this so?

3-Dimensional Structures of D-Tartaric Acid (left) and L-tartaric acid (right)



Each ball is an atom. Black balls are carbon, pink is hydrogen, and red is oxygen. Each structure has the same chemical structure but is arranged differently in 3-dimensional space. As such, they have different chemical properties.

Upon examination under a microscope, he found that every crystal of tartaric acid looked exactly like the other, but that the crystals of paratartaric acid were a 1:1 mixture of two types. Each crystal looked the same, but was the polar opposite of the other, like holding up a mirror to one of them. He then painstakingly hand-separated each crystal under a microscope using tweezers and placed each crystal in a solution. He referred to the configurations as stereoisomers of each other. When he shone polarized light through them, he found that one solution rotated light to the right, while the other rotated light to the left. He called the crystal that rotated light to the right the D-stereoisomer (for dextro, Latin for 'right'), whereas the crystal that rotated light to the left was called the L-stereoisomer (for levo, Latin for 'left'). Each stereoisomer was said to be an enantiomer, a particular type of stereoisomer that is a non-superimposable mirror image of each other. A 1:1 mixture of L- and D-enantiomers is called a racemic mixture. Pasteur found that tartaric acid rotated light to the left because it was a pure L-enantiomer and that paratartaric acid was tartaric acid, but was a racemic mixture of the two enantiomers that did not rotate polarized light.

Compounds like tartaric acid are said to be chiral, which is Latin for 'hand.'. Pasteur's discovery initiated the field of stereochemistry, which examines the arrangement of chemicals in three-dimensional space. It is now known that chiral compounds have an atom, usually carbon or sulfur, with four different atoms attached to it. Surprisingly, enantiomers exhibit different chemical and physical properties. Compounds without chirality are achiral, having superimposable mirror images, and have the same physical and chemical properties.

If you think about it, enzymes and proteins have specific 3-dimensional configurations. Every protein in the body is composed of amino acids. All those amino acids are chiral, having an L-

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

How Important is Chirality?

configuration. Why evolution chose the L-configuration over the D-configuration is one of the great mysteries of biology. All of those amino acids being locked in an L-configuration have important implications for the protein's structure and activity.

Drugs can also be chiral, and as such, their enantiomers have different chemical properties. For a drug to bind to a protein's active site, it must have the proper configuration. If it does not, the drug may not work at all. To make matters more confusing, biologists use D- and L-notation for carbohydrates and amino acids, but often use a different notation for drugs: the S/R notation. Both notations refer to handedness, and while they are not exactly the same, R- is generally used for right-handedness and S- for left-handedness. But a D-configuration is not always the same as an S-configuration, and an L-configuration is not always the same as an S-configuration. Needless to say, it's complicated.

Going back to the main question, how important is chirality? Let's ask thalidomide. Thalidomide, which was introduced earlier in the book, was used as a treatment for morning sickness in the 1950s and 1960s. It was found that thalidomide caused birth defects and, as a result, many of those children who were born had shortened or absent limbs. It turns out that thalidomide is chiral with a single chiral center. In 1979, it was reported that only the S-enantiomer is responsible for thalidomide's teratogenicity [148]. The R-enantiomer causes sedation. The authors of that study suggested that the entire tragedy could have been avoided if they had used the pure enantiomer as the drug product.

While not as notable as thalidomide, praziquantel is used to treat schistosomiasis, a parasitic disease caused by flatworms primarily found in drinking water, at a dose of 20 to 25 mg/kg three times daily. For a 70 kg person, the dose requirement is about 1400 mg three times daily. For decades, praziquantel was formulated for adults as a 600 mg tablet. For an adult, 1 to 3 tablets three times a day is sufficient. However, for a child, swallowing a 600 mg tablet is challenging due to its size. To make it easier to swallow, tablets are often crushed into a consistency like applesauce and then eaten. But even doing that, kids hate it because it tastes terrible. Praziquantel is chiral, and it turns out that the S-enantiomer is entirely responsible for its taste. The R-enantiomer is responsible for its anti-parasitic activity. In 2012, the nonprofit Praziquantel Pediatric Consortium was launched to develop a palatable pediatric formulation of praziquantel, utilizing only the R-enantiomer. Arpraziquantel was approved for use by the European Medicines Agency in 2023 for the treatment of schistosomiasis in preschool children 3 months to 6 years.

Receptor Occupancy and Drug Efficacy

It is generally accepted that a drug's effect increases with increasing dose until a maximum is reached. This principle follows from the interaction between a drug and its receptor. A drug (D) binds to its receptor (R), forming a drug-receptor complex (DR) that, after some period of time, dissociates back to its components D and R. This can be written as:

$$\text{Drug} + \text{Receptor} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{Drug-Receptor Complex} \tag{20}$$

The rate of the forward reaction, which forms the Drug-Receptor Complex, depends on the concentration of the drug and receptor, called [D] and [R], respectively, and is equal to $k_{on}[D][R]$, where k_{on} is called the rate constant for the binding action. The reverse reaction, where the Drug-Receptor Complex dissociates into drug and receptor, depends on the concentration of the complex, called [DR], and is equal to $k_{off}[DR]$, where k_{off} is also a rate constant. After dissociation, D and R remain the same as before; they can bind again to form DR if they are in proximity with each other.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

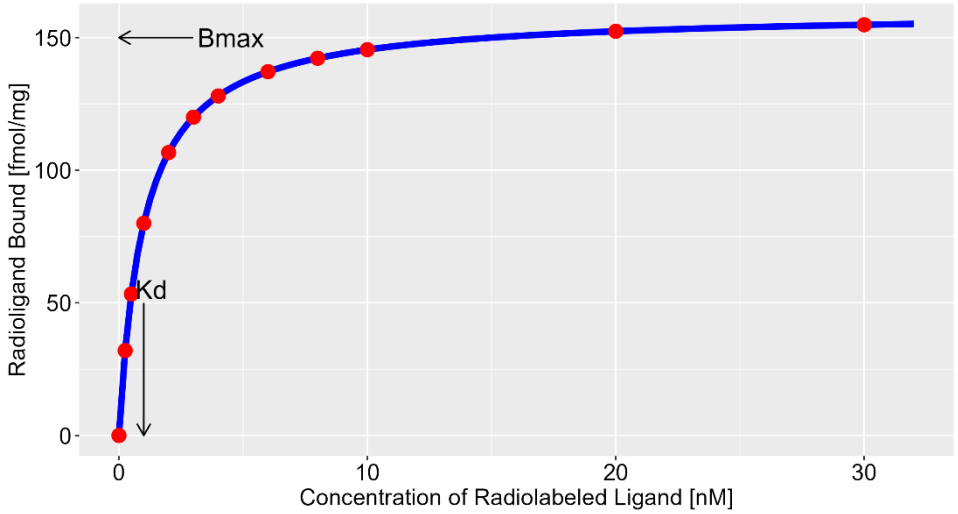


Figure 45. Saturation binding experiment to a receptor. As the concentration of radiolabeled ligand increases, the amount bound increases until a limit is reached. This is Bmax. The radioligand concentration that leads to 50% of Bmax equals the K_d.

At equilibrium, based on the law of mass action, the association rate equals the dissociation rate, i.e., the two rates are equal. Hence,

$$k_{on}[D][R] = k_{off}[DR]. \tag{21}$$

By rearrangement,

$$\frac{k_{off}}{k_{on}} = \frac{[D][R]}{[DR]} = K_d \tag{22}$$

where K_d is the dissociation constant for the reaction, measuring the affinity between the enzyme and substrate. K_{on} is a measure of the strength of the association between a ligand (drug) and its receptor and K_{off} is a measure of the strength of the dissociation between a ligand (drug) and its receptor, which is not to be confused with K_d. A small K_d indicates tight receptor binding, while a large K_d value indicates weak binding. Note that K_d and the rate constants k_{on} and k_{off} are different. Consider the case where the drug occupies half the receptors, and the other half are unoccupied. This occurs when [D] = [DR] and only happens when [D] = K_d. Hence, K_d is the drug concentration at equilibrium, where half the receptors are occupied with the drug.

Another concept of interest is receptor occupancy. The fractional occupancy, i.e., the proportion of receptors occupied by the drug, is equal to

$$\text{Fractional Occupancy} = \frac{[DR]}{[R]+[DR]}. \tag{23}$$

Using the definition for K_d in Eq. (22) and rearrangement into Eq. (23),

$$[R] = \frac{K_d[DR]}{[D]}. \tag{24}$$

Putting Eq. (24) into Eq. (23) and expressed as a percent then gives

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Rate Theory: Example of an Alternative to Occupancy Theory

Rate theory, proposed by William Paton in 1961, states that the drug response is proportional to the rate at which the drug-receptor complex is formed. Whether a drug is an agonist, antagonist, or partial agonist is determined by how long the drug exists in the drug-receptor complex [DR] state. With each association of drug and receptor, there is a "quantum" of excitation at the receptor. The more often a drug associates, dissociates, and then reassociates, the greater the response of the drug. For this to occur, drugs must dissociate from their receptor rapidly. Hence, the dissociation rate constant determines whether a drug is an agonist, partial agonist, or antagonist under this model. While useful in explaining many observations, such as the persistent effect of antagonists and the short duration of agonist activity, it had its limitations. For example, the theory predicts that antagonists will exhibit some residual agonist activity because, over time, they dissociate from the drug-receptor complex. While this has been seen with some antagonists, it does not occur with most. Today, rate theory, while consistent with many observations, is considered of limited usefulness.

$$\text{Percent Occupancy} = \frac{[\text{D}]}{[\text{D}] + \text{K}_\text{d}} \times 100\% \,. \tag{25}$$

From Eq. (25) , it can be seen that when drug concentration [D] equals zero, occupancy equals zero. When [D] equals K_d, the occupancy equals 50%, and when [D] >> K_d, occupancy $\rightarrow$ 100%. An alternative way to write Eq. (25) is

$$[\text{DR}] = \frac{\text{Bmax} \times [\text{D}]}{[\text{D}] + \text{K}_\text{d}} \tag{26}$$

where Bmax is the maximal amount of drug that can bind to a receptor. This is illustrated in Figure 45. Although some assumptions were made in deriving this equation ([R] and [DR] are at equilibrium, binding to the receptor is reversible, D is not changed upon dissociation from DR, etc.), it has been found useful for explaining the behavior of many drugs.

One of the earliest theories of drug response, proposed by Alfred Clark in 1926, is that the intensity of response is proportional to the number of occupied receptors and persists as long as the drug occupies them. This can be written mathematically as:

$$[D] + [R] \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [DR] \Rightarrow efficacy \,. \tag{27}$$

This is known as the *receptor occupation theory of agonism [149]*. Two major assumptions of this model are that the maximal effect occurs when all receptors are completely occupied and that all drugs elicit the same effect when receptors are maximally occupied. This latter prediction was shown to be incorrect by Everhardus Ariens in the 1950s, who demonstrated using a series of chemical analogs that some analogs could similarly maximally occupy their receptors but did not demonstrate equal effect [150]. Some analogs showed greater effect than others. This led Ariens to propose the concept of partial agonists and intrinsic activity. Partial agonists are drugs that can bind to a receptor maximally but exhibit a lower maximal response than other agonists at the same receptor. Intrinsic activity is the ability of a drug to produce maximal responses.

Robert Stephenson expanded on Ariens' concepts in 1954. He proposed that the response was some unknown nonlinear positive function of receptor occupancy and did not need to be proportional to receptor occupancy [150]. He further proposed that maximal occupancy was not needed for maximal effect – that the maximal effect could occur if only a small percentage of receptors were occupied.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

This idea was not proven until years later by Robert Furchgott, who later won the Nobel Prize for his work, when he demonstrated that cells indeed had "spare receptors" beyond what was needed for maximal effect. Based on Stephenson's ideas, Ariens expanded the concept to include the idea of a stimulus threshold – an effect may only occur when a minimum number of receptors are stimulated. In some cases, a significant proportion of receptors must be occupied before an effect is observed.

Affinity and Intrinsic Activity

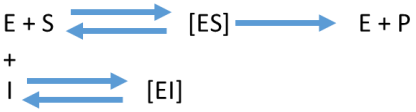
Agonists are ligands that bind to receptors and elicit an effect. They can be natural ligands, such as neurotransmitters, or synthetic drugs. Agonists that bind to receptors are defined by their affinity and intrinsic activity. Affinity is equal to K_d and measures how tightly the binding of an agonist is to a receptor. Intrinsic activity measures how likely the agonist is to activate a receptor when it is bound to the receptor.

With the discovery of GCPRs and the advent of molecular biology, occupation theory has undergone modifications (e.g., the multistate receptor model, where the receptor exists in multiple active and inactive states) to explain recent findings, but the basic concepts remain intact. Since occupancy theory does not explain every observation, it has given rise to competing theories, such as rate theory, macromolecular perturbation theory, and drug-target residence time theory. However, occupation theory explains most situations and has the widest acceptance and broadest application.

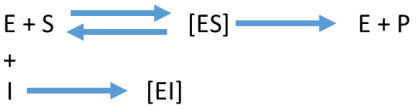
Enzymes and Transporters

Enzymes and transporters are unique in that the concept of an agonist does not apply. We do not develop drugs to turn on an enzyme or a transporter; we develop antagonists to block them. Enzymes catalyze chemical reactions by lowering the energy requirements for the reaction to occur. Consider the reaction of substrate S being converted to product P. Under classical enzyme theory, S reacts with the enzyme E, forming a substrate-enzyme complex [ES], which then converts to product P and releases the enzyme, allowing it to repeat the process. This is illustrated as follows:

There are different types of antagonist drugs, but the most common is a competitive antagonist. Competitive antagonists (denoted as I) compete with the enzyme, reducing the available enzyme for the conversion of S to P, as follows:



The same scheme applies to transporters, where S is transported across the cell membrane instead of being converted to P. There are other types of antagonists, but an interesting one is suicide inhibitors, like aspirin and disulfiram, in which the inhibitor irreversibly binds to the enzyme, and the reverse reaction from [EI] back to I and E does not occur:



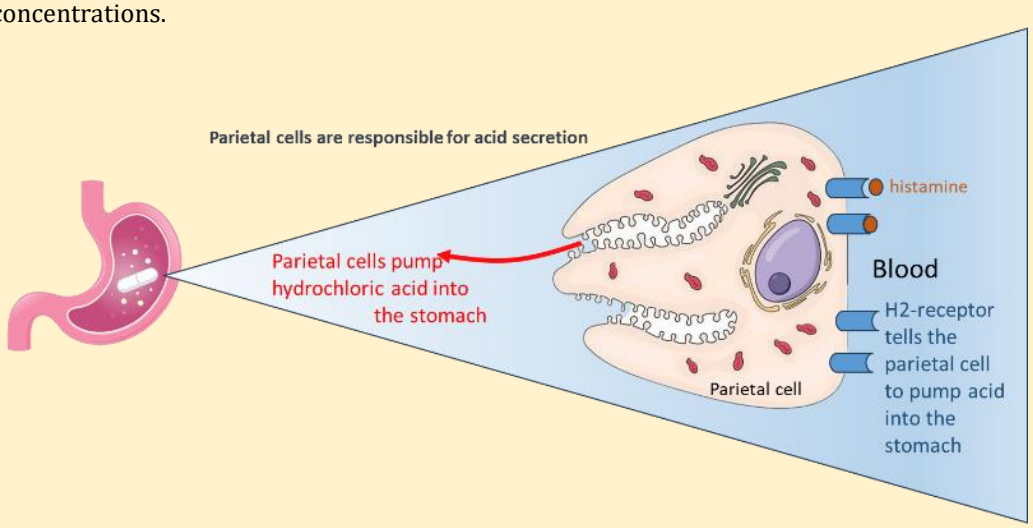
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Omeprazole and Famotidine Are Two Drugs Used to Treat Gastroesophageal Reflux (GERD)

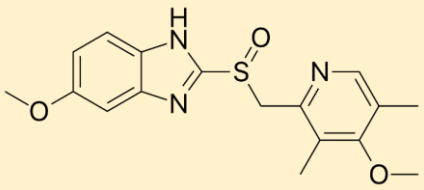
Omeprazole and famotidine are two drugs used to treat gastroesophageal reflux disease (GERD), a condition in which acid flows back from the stomach into the esophagus, causing heartburn. If untreated, GERD can lead to esophageal ulcers and an increased risk of esophageal cancer. Both drugs, however, work by different mechanisms.

Omeprazole works by irreversibly binding to the H+/K+-ATPase in parietal cells in the stomach, thereby inhibiting acid secretion and increasing stomach pH, providing relief for an acid stomach. However, it takes days for omeprazole to have its full effect, and it is not a good drug for quick relief. Because omeprazole irreversibly binds to its receptor, once drug administration stops, the stomach only recovers after making new ATPase, which also takes days and provides longer-lasting relief.

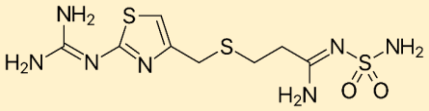
In contrast, H2-blocking drugs, like famotidine or cimetidine, bind to the H2-histamine receptor located on the blood side of parietal cells, but do so competitively. As drug concentrations decline, the degree of acid inhibition declines as well. However, the onset of the effect is rapid. In other words, the onset and offset of the effect parallel drug concentrations.



Omprezole



Famotidine



Structures courtesy of Wikipedia.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

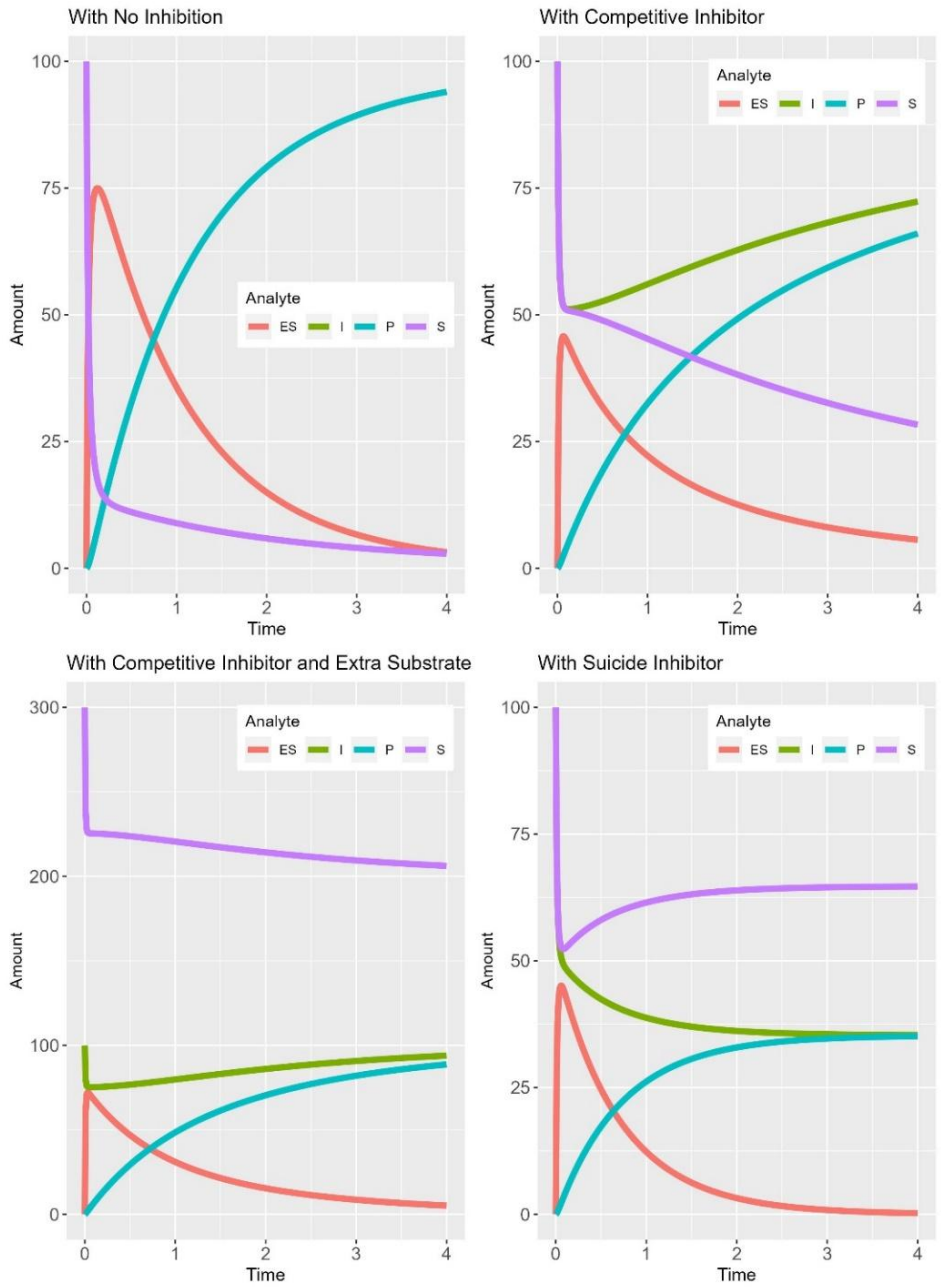


Figure 46. Illustration of the reaction kinetics for a substrate S in the absence (top left) and presence of a competitive inhibitor with equal affinity as the substrate (top right). The bottom left plot shows the reaction kinetics in the presence of a competitive inhibitor and a much higher initial substrate concentration. Competitive inhibition can be overcome by adding an excess of substrate. The bottom right plot shows a suicide inhibitor whose backward reaction from [EI] does not occur.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

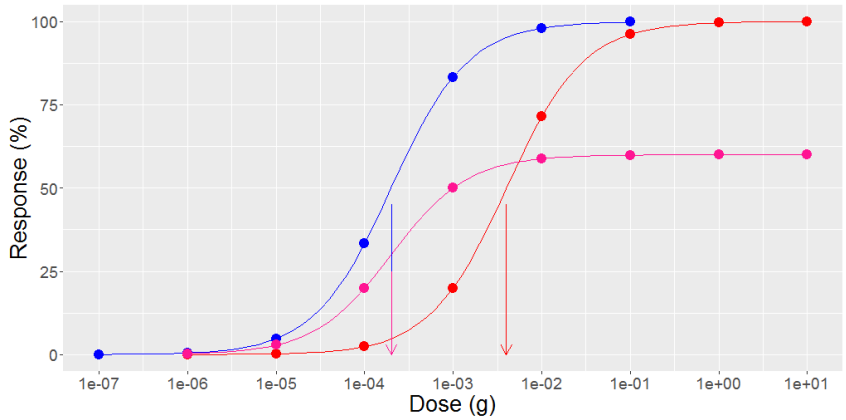


Figure 47. Dose-response profile for three different drugs. The red and blue lines are complete or full agonists, but the pink line is a partial agonist because its maximal response is only a fraction of the full agonist. The arrows are the ED_{50}, the dose at which 50% of the maximal response occurs. The ED_{50} for the blue and pink drugs is the same (which is why the arrow has two colors). Despite the blue and red drugs having the same efficacy, the blue drug is more potent, as its ED50 is significantly lower. The pink drug is not as efficacious as the blue drug, but the blue and pink drugs are equally potent because they have the same ED_{50}.

Returning to competitive inhibition, suppose we now develop a drug to block an enzyme or a transporter. How does this affect the conversion of S to P or the transport of S? Without any math, it's easy to see that the rate of transport or rate of conversion to P will decrease. This is illustrated in Figure 46. The turquoise line is the amount of product P formed over time. Without an inhibitor, at time 4, the amount of product formed is almost 100% of the initial substrate amount (top left plot). But when a competitive inhibitor of equal affinity to the enzyme as the substrate is added to the system (top right plot), less than 70% of the substrate is converted to the product at time 4. The competitive inhibitor slows the rate of conversion. However, the key aspect of competitive inhibitors is that they can be overcome by adding more substrate to the system. When we add three times the initial substrate amount to the system, the amount of product formed at time 4 is almost the same as in the absence of the inhibitor (bottom left).

For comparison, a suicide inhibitor was added under the same conditions (bottom right), showing that even more product inhibition occurs than with a competitive inhibitor. Hence, by binding to the enzyme, the drug destroys the target it was meant to affect. Only by synthesizing a new enzyme or transporter does the function return. Examples of irreversible drug effects are proton pump inhibitors (PPIs), aspirin, warfarin, and some anticancer drugs. For example, PPIs irreversibly bind to H+/K+-ATPase receptors in gastric cells in the stomach, whose normal function is to secrete hydrogen ions into the stomach, thereby aiding food digestion. PPIs increase stomach pH but take days to reach their full effect. Hence, if one has an acid stomach, a PPI is not the drug of choice for fast relief; it is, however, the drug of choice for long-acting relief if the drug is continually taken. Even though the elimination half-life of a drug like omeprazole is only an hour, because the turnover rate for synthesizing new H+/K+ ATPase is 24 hours, omeprazole can be given once daily for maximal effect.

The Dose-Response Relationship

The dose-response relationship is the foundation of pharmacology, describing the relationship between the response of a drug and its dose. When drugs are given to animals or humans, the effects at the receptor translate to an observed effect. This is the [DR] $\Rightarrow$ response step in Eq. (27). For example, salivation occurs when M1 and M3 muscarinic receptors are stimulated with pilocarpine, and dry mouth occurs when muscarinic antagonists, such as scopolamine, are administered. Hence,

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

when the administered dose is plotted against response, a profile similar to that of receptor occupancy is observed. This time, however, the salient features of the dose-response profile are its efficacy and ED_{50}. Efficacy is the maximal observed response, whereas ED_{50} is the dose that produces 50% of the maximal response. These concepts are illustrated in Figure 47 for three different drugs: two full agonists and a partial agonist.

Antagonists, as described earlier, block the effect of agonists and, on their own, have no effect. Antagonists affect the dose-response profile, and how the profile is altered defines the type of antagonist a drug is. Competitive antagonists compete with an agonist at the receptor's active site, shifting the dose-response profile to the right. As the antagonist dose increases, the profile shifts further to the right. The effect of a competitive antagonist can be overcome by increasing the agonist dose. An example of competitive antagonism is the effect of propranolol on beta-adrenergic receptors in response to isoproterenol.

A noncompetitive antagonist either binds irreversibly to the active site so that the agonist cannot bind to the receptor or binds to an allosteric site on the receptor that reduces the binding of the agonist to the receptor. Noncompetitive antagonists do not affect the agonist's ED50 but decrease the agonist's maximal effect. The effects of a non-competitive antagonist cannot be overcome by increasing the agonist dose. These concepts can be seen in Figure 47. This time, however, think of the blue line as the agonist, the red line as the agonist in the presence of a competitive antagonist, and the pink line as the agonist in the presence of a noncompetitive antagonist. An example of noncompetitive antagonism is the effect of phenoxybenzamine on norepinephrine at alpha-adrenergic receptors.

Kinetics of Drug Response – Single Dose

Until now, the effect of time has been ignored in developing the relationship between drug concentration, receptor occupancy, and response. This is perfectly acceptable in *in vitro* systems where the drug bathes the receptor, such as occurs in a test tube, and drug concentrations remain constant. However, in animals, time cannot be ignored, as pharmacokinetics also plays a significant role. Drug concentrations do not remain constant in an *in vivo* system; they change over time, and as a result, receptor occupancy and effects also change.

Regarding the effect of time on drug response, drugs can be classified as either rapid equilibrium or slow equilibrium. For an agonist drug that has a rapid equilibrium between the plasma compartment and the site of action, peak drug concentrations occur at the same time as the peak drug effect (left column of Figure 49). When drug concentrations are in rapid equilibrium with the target, plotting drug concentrations against the effect results in a hyperbolic plot, such as that seen with receptor occupancy. Taking a closer look, the two central parameters of this plot are E_{max}, the maximal effect, and EC_{50}, the concentration that produces 50% of the maximal effect. This profile can be modeled mathematically as

$$\text{Effect} = \frac{E_{max} \times [C]}{EC_{50} + [C]}. \tag{28}$$

There are two limiting cases to this model. When drug concentration [C] is much greater than EC_{50}, then Effect $\rightarrow$ Emax. In other words, the maximal effect will be observed when drug concentrations are very large. At the other extreme, when drug concentrations are less than the EC50, the effect is directly proportional to the drug concentration. As shown in Figure 48, whatever drug concentrations are, whether they are at the high or at the low end, their location on the concentration-effect profile will also determine what happens to the effect when drug concentrations change.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

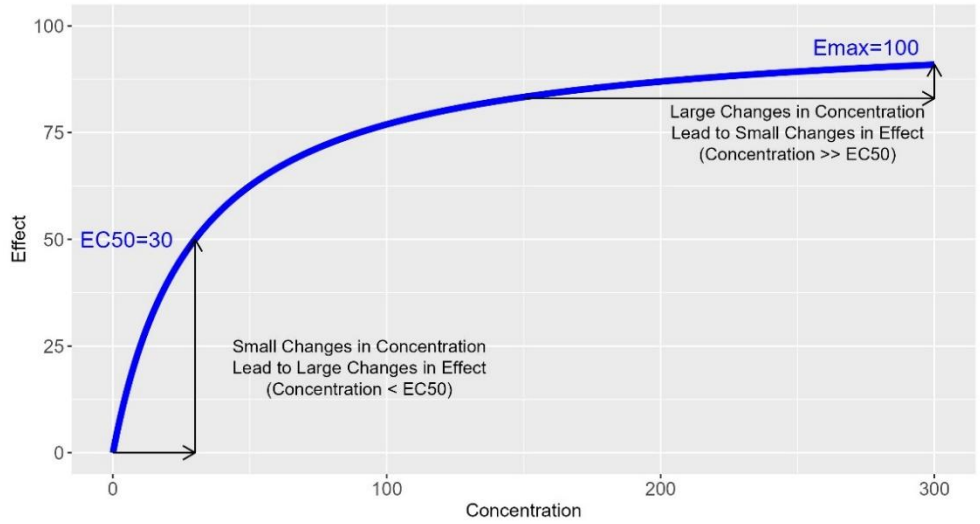


Figure 48. Result of changes in drug concentration on effect. The result of a change in drug concentration will depend on where you are on the concentration-effect profile. When concentrations are less than the EC_{50}, small changes in drug concentration lead to large changes in effect, but when drug concentrations exceed the EC_{50}, large changes in drug concentration led to small changes in effect.

For a slowly equilibrating drug, however, there is a delay between the peak drug concentration and peak effect (right column of Figure 49). A counter-clockwise hysteresis loop is observed when concentration is plotted against the effect. Notice what happens when the drug concentration is 75. A concentration of 75 is observed twice in an oral drug concentration-time profile, once at ~1 time unit after dosing (when concentrations are rising) and again at ~9 time units after dosing (when concentrations are declining). The effect is the same for a rapidly equilibrating drug, ~71, regardless of whether concentrations increase or decrease. However, the effect depends on time for a slowly equilibrating drug. At 1-time unit after dosing, the effect is ~1.5, but at 9-time units after dosing, at the same concentration, the effect is ~21. The effect is greater at 9-time units after dosing compared to 1-time unit after dosing, even though concentrations are declining in this part of the concentration-time profile.

Consider this scenario now. You have a severe headache, and you want to take some medicine for it. The label says to take one pill every six hours. You decide that this headache is really bad and that you need two pills. What will happen to the effect? Will the effect be stronger? You've doubled the dose, will you double the effect? Surely, the effect will last longer, right? Will the onset be faster? The concentration-effect profile allows us to answer these questions. Figure 50 shows what happens when the dose is doubled for three different EC_{50} values. In all cases, dose doubling resulted in a greater effect, but only when drug concentrations were below the EC50 did it approximate a doubling of the effect (see the bottom row of plots). In general, dose doubling can result in a slightly quicker onset of effect because therapeutic concentrations at the site of action are achieved more quickly. Still, in this example, the difference in onset was negligible.

In all 3 cases, dose doubling resulted in a longer effect. Why? Suppose a 20% effect is needed for headache relief (which is the black line in the plots on the right). When the dose was doubled, and drug concentrations exceeded the EC_{50}, the effect was prolonged by about the same amount in all three cases. This had to do with the half-life of the drug because it turns out that the effect increases by one half-life for every doubling of the dose. In this case, the half-life of the drug was 7 hours, and if you look carefully at Figure 50 you will see that in every case the effect was prolonged by about 7 units. Doubling the dose increases the effect time of the drug by 1-half-life.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

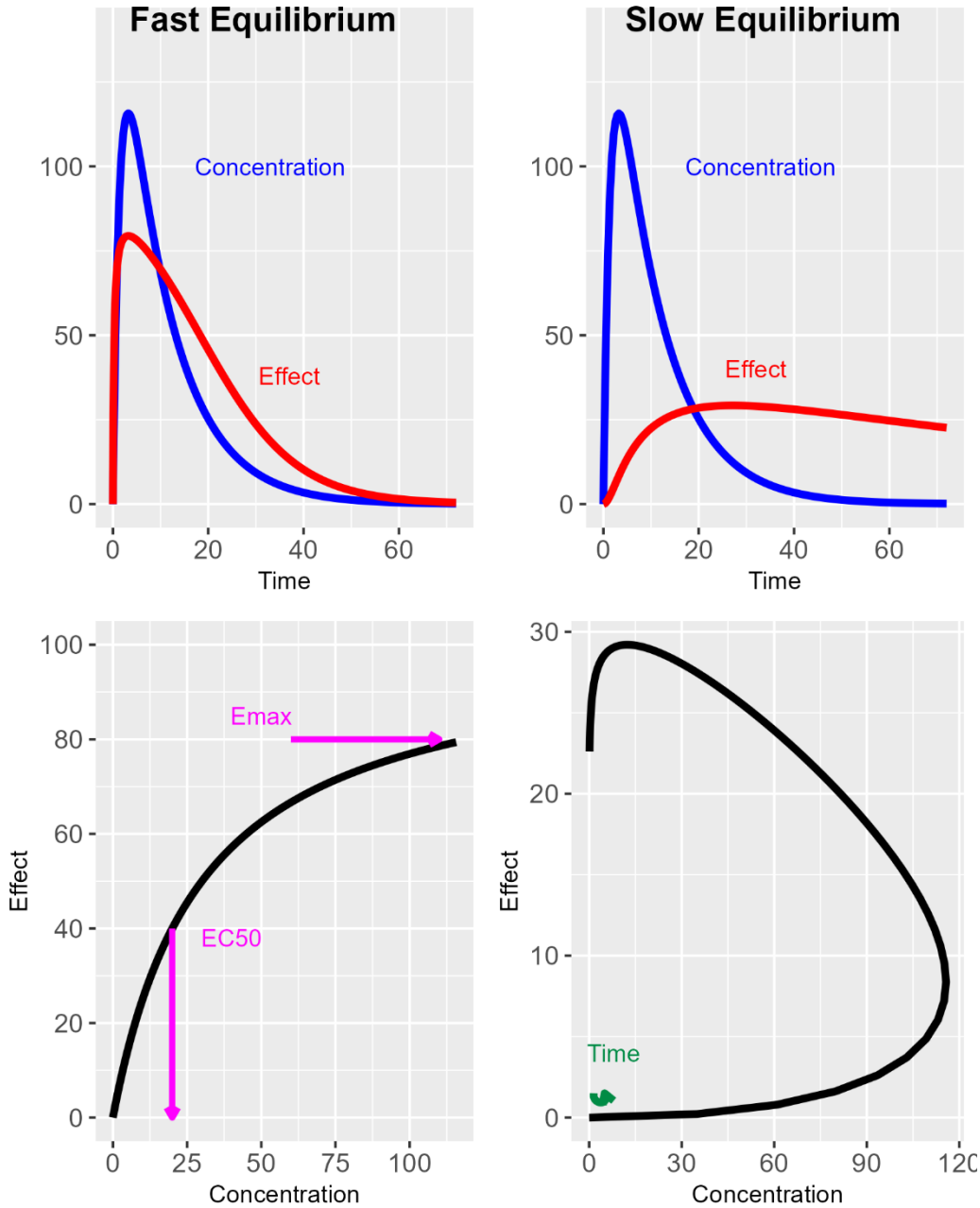


Figure 49. Concentration- and effect-time plots (top row) and concentration-effect plots (bottom row) for fast (left column) and slowly (right column) equilibrating drugs. Drugs that equilibrate quickly with their receptor have a sigmoid or hyperbolic shape (bottom left), whereas drugs that are slowly equilibrating show a counter-clockwise hysteresis loop (bottom right).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

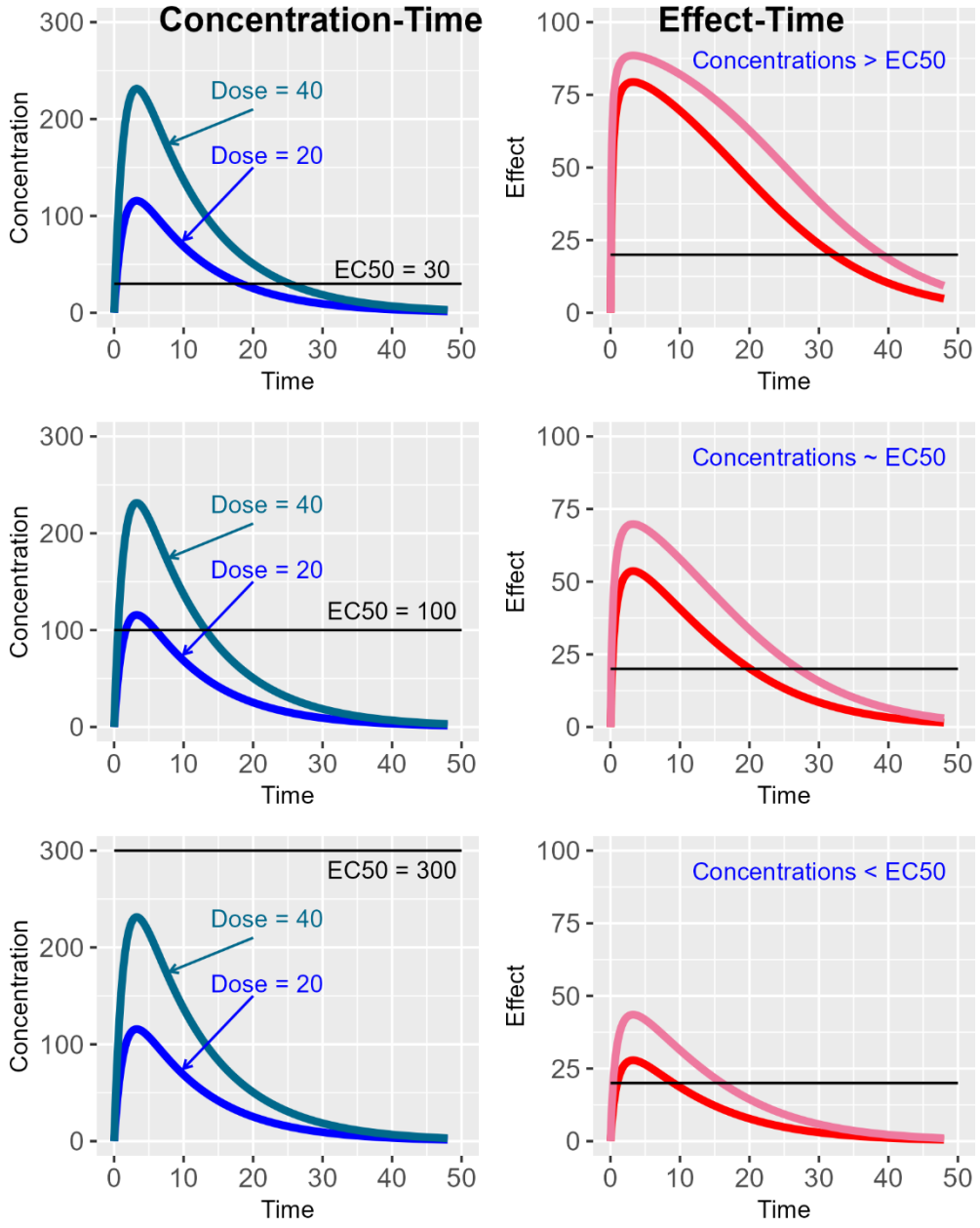


Figure 50. Concentration-time (left column) and effect-time (right column) profiles for a rapid-equilibrium agonist drug when the administered dose is doubled, stratified by EC50 value. Only when drug concentrations are much lower than the EC50 does dose doubling result in a doubling of the peak drug effect. In all cases, the time to peak drug effect remains the same. Because the half-life of effect follows the half-life of the drug, dose doubling results in a prolongation of effect (it turns out to increase the effect by one half-life) but does not double the length of the effect.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

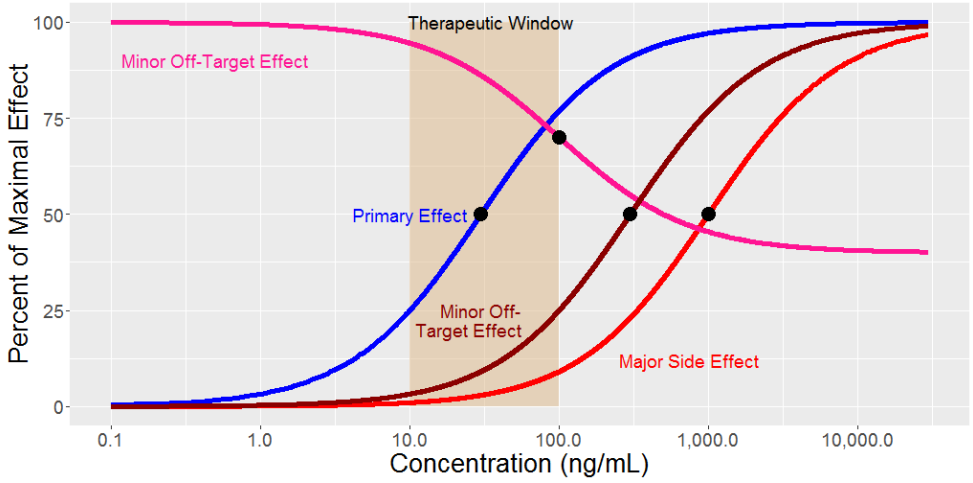


Figure 51. Concentration-effect profile for a drug for four different effects – one primary effect and three side-effects, one of which is antagonism (pink line). Each effect has a unique EC_{50} or IC_{50}, shown as black dots on the curve. The therapeutic window for the drug is shown by the light brown-shaded area (10–100 ng/mL).

For a slowly equilibrating drug, however, there is a delay between the peak drug concentration and peak effect (right column of Figure 49). A counter-clockwise hysteresis loop is observed when concentration is plotted against the effect. Notice what happens when the drug concentration is 75. A concentration of 75 is observed twice in an oral drug concentration-time profile, once at ~1 time unit after dosing (when concentrations are rising) and again at ~9 time units after dosing (when concentrations are declining). The effect is the same for a rapidly equilibrating drug, ~71, regardless of whether concentrations increase or decrease. However, the effect depends on time for a slowly equilibrating drug. At 1-time unit after dosing, the effect is ~1.5, but at 9-time units after dosing, at the same concentration, the effect is ~21. The effect is greater at 9-time units after dosing compared to 1-time unit after dosing, even though concentrations are declining in this part of the concentration-time profile.

Consider this scenario now. You have a severe headache, and you want to take some medicine for it. The label says to take one pill every six hours. You decide that this headache is really bad and that you need two pills. What will happen to the effect? Will the effect be stronger? You've doubled the dose, will you double the effect? Surely, the effect will last longer, right? Will the onset be faster? The concentration-effect profile allows us to answer these questions. Figure 50 shows what happens when the dose is doubled for three different EC_{50} values. In all cases, dose doubling resulted in a greater effect, but only when drug concentrations were below the EC50 did it approximate a doubling of the effect (see the bottom row of plots). In general, dose doubling can result in a slightly quicker onset of effect because therapeutic concentrations at the site of action are achieved more quickly. Still, in this example, the difference in onset was negligible.

In all 3 cases, dose doubling resulted in a longer effect. Why? Suppose a 20% effect is needed for headache relief (which is the black line in the plots on the right). When the dose was doubled, and drug concentrations exceeded the EC_{50}, the effect was prolonged by about the same amount in all three cases. This had to do with the half-life of the drug because it turns out that the effect increases by one half-life for every doubling of the dose. In this case, the half-life of the drug was 7 hours, and if you look carefully at Figure 50 you will see that in every case the effect was prolonged by about 7 units. Doubling the dose increases the effect time of the drug by 1-half-life.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

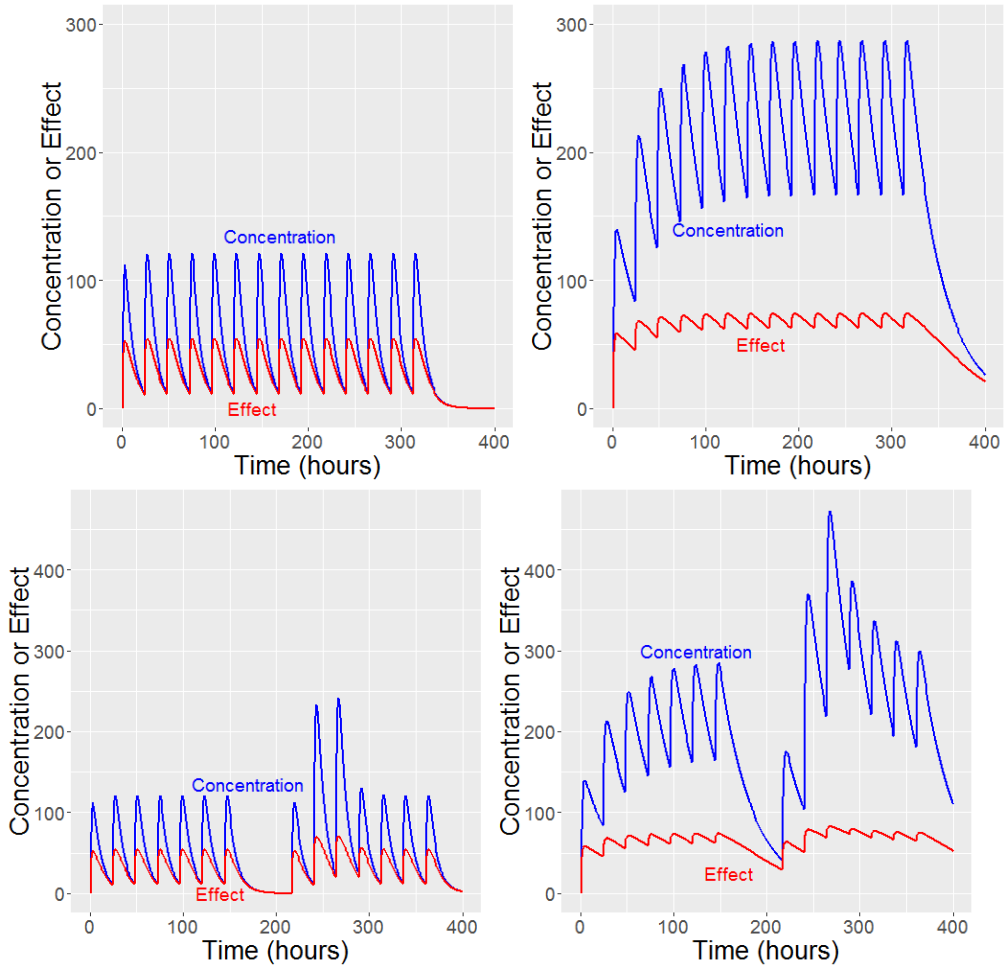


Figure 52. Concentration-time (blue) and effect-time (red) profiles for a rapid-equilibrium drug with an elimination half-life of 6 hours (left plots) or 24 hours (right plots). All other aspects of the drugs are the same. The top plots show the profiles when compliance is 100% - the drug is taken every 24 hours. The bottom plots show the profiles for non-100 % compliance. Sometimes a dose is missed, sometimes twice the dose is taken. Notice that for a short half-life drug, the effect profiles are similar to the concentration-time profiles. There is a significant difference between the peak and the trough effects. With missed doses, the effect is not observable. With extra doses, there is little increase in effect. With a drug having a longer half-life, there is accumulation of the effect, much like the accumulation seen with the concentration-time profile. Furthermore, the difference between peak effects and trough effects is smaller; the effect of missed doses is less severe, and the effect of extra doses is not particularly noticeable. When the drug effect is relatively unchanged in the presence of poor compliance, the drug is said to be 'forgiving' with respect to compliance.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

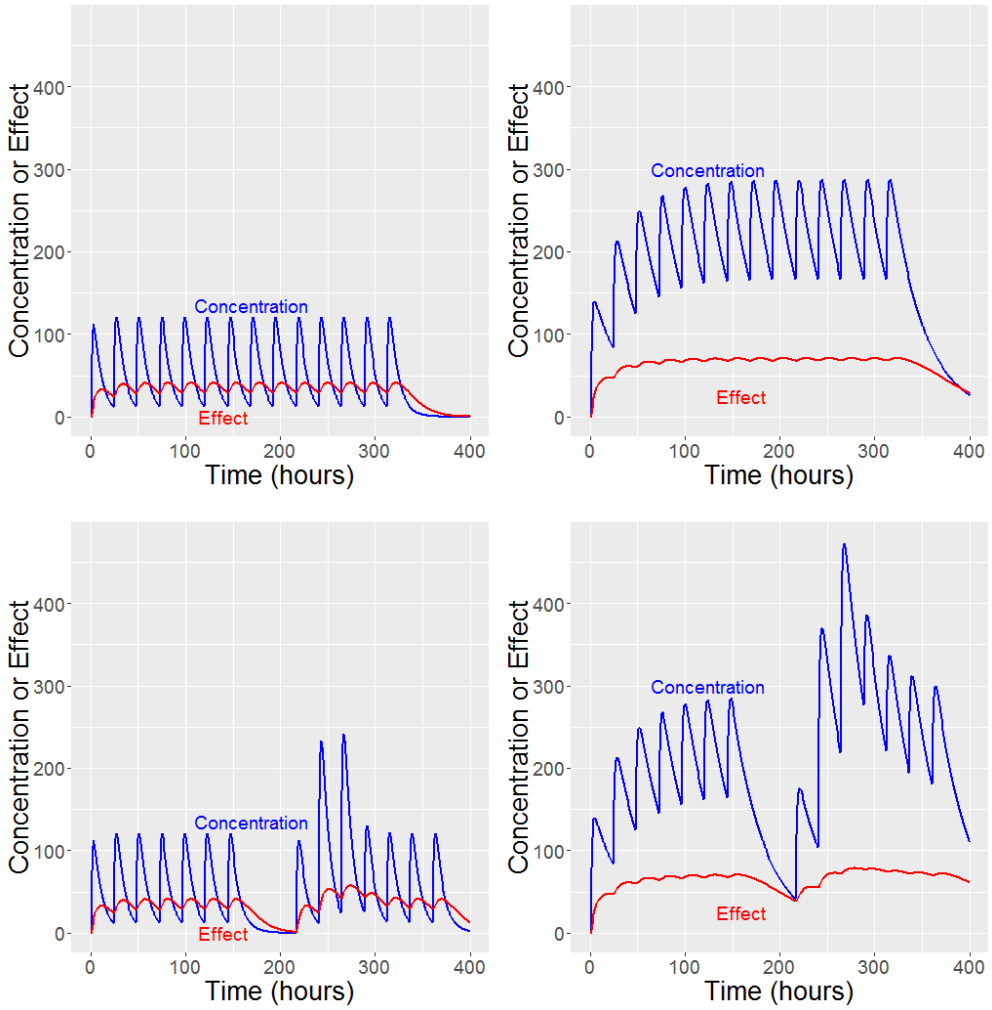


Figure 53. Concentration-time (blue) and effect-time (red) profiles for a slow-equilibrium drug with an elimination half-life of 6 hours (left plots) or 24 hours (right plots). The plots have the same parameters as in Figure 52, except that there is a 12-hour delay. The top plots show the profiles when compliance is 100% - the drug is taken every 24 hours. The bottom plots show the profiles for non-100 % compliance. Sometimes a dose is missed, sometimes twice the dose is taken. Notice that for the short half-life drug, there is a buffering seen with the drug effect. The drug effect does not show the same peak-to-trough variability as a rapid-equilibrium drug. With a long half-life drug, the drug effect is almost constant. This is because the delay buffers the link between drug concentration and effect. Drug effect is even less affected by compliance in both cases, making the drug even more forgiving than with a rapid equilibrium drug.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

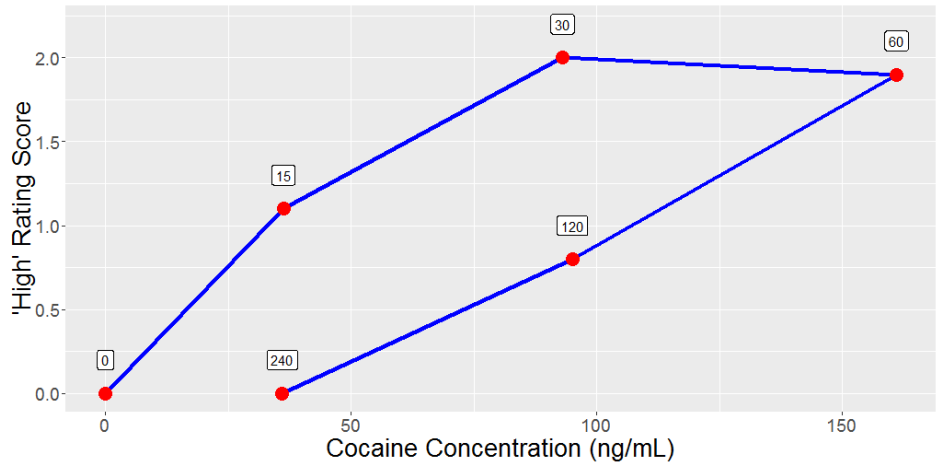


Figure 54. Concentration-effect plot of "high" after intranasal administration of cocaine in 4 volunteers. Notice that in contrast to the counter-clockwise hysteresis seen between concentration and effect with slowly equilibrating drugs (Figure 49), a clockwise hysteresis is observed. Cocaine concentrations at 15 min and 240 min after administration are similar. Still, the effect is different – the 'high' seen at 240 min after administration is practically zero compared to 15 min after dosing, despite having similar concentrations. This is evidence of tachyphylaxis. Data are redrawn from [151].

Most drugs have multiple effects, and each effect has its own concentration-effect profile (Figure 51). Some of these effects are therapeutic; some are not. Of those that are not, some may be minor side effects, and some may be major side effects that could lead to severe debilitation or death. The ratio of the EC_{50} of the severe side effect to the EC_{50} of the therapeutic effect is known as the therapeutic index. The larger the therapeutic index, the safer the drug. In the case of Figure 51, the EC_{50} of the major side effect was 1000 ng/mL, while the EC_{50} of the primary effect was 30 ng/mL. Hence, for this drug, the therapeutic index was 33, which makes this drug quite safe. Even if a drug interaction increases drug concentrations 10-fold (which is large), there is still a 3-fold safety margin. To put this in perspective, alcohol has a therapeutic index of 10, but digoxin only has a therapeutic index of 2 (which is narrow). Related to the therapeutic index is the therapeutic window, shown by the light brown shaded area in Figure 51. The therapeutic window is the range of drug concentrations that are effective without toxicity. The wider the window, the safer the drug. In this example, the therapeutic window, defined by the occurrence of a major side effect, is 10-100 ng/mL. Sometimes, you will see the therapeutic window and index expressed as a ratio of the dose that produces toxicity to the safe dose.

Kinetics of Drug Response – Multiple Dosing

Most drugs are not taken only once. Most drugs are taken repeatedly, usually once a day, sometimes two or three times a day, for some period of time, which may be indefinite. The effect-time profile of multiple dosing depends on the drug's half-life, whether the drug is in rapid or slow equilibrium, and the patient's compliance with the dosing regimen. Compliance, sometimes called adherence, is whether a patient takes the drug as prescribed and how consistently they take the drug. For instance, patients may miss a dose or accidentally take twice the intended dose (forgetting they had already taken their dose that day). Furthermore, patients may not take the drug at the same time every day. If a patient takes a drug before bedtime, they could take the drug anywhere from early evening to late evening. The effect of compliance on drug effect is complex.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Figure 52 presents the concentration- and effect-time profiles for a rapid-equilibrium drug with a short (6 hours) or long (24 hours) elimination half-life, and whether the patient is compliant or noncompliant. For short-half-life drugs, the effect-time profile resembles a series of single-dose effect-time profiles because there is no drug accumulation, and drug concentrations approach zero before the next dose is administered. When a dose is missed, it is problematic for the patient because the drug does not take effect until the next dose. Doubling the dose accidentally has a small increase in maximal effect and an increase in drug effect by one half-life.

Drug concentrations accumulate for a drug with a long half-life, and the concentration-time profile shows less peak-to-trough variability than with single dosing. Hence, the drug effect will also show less variability. Additionally, as drug concentrations increase, the effect approaches its maximum. The effect in Figure 53 is almost flat. Missing a dose has almost no effect because concentrations decline by only two half-lives before the next dose is administered. Furthermore, doubling the dose has almost no effect. With long half-life drugs that rapidly equilibrate, the effect is buffered with multiple dosing. For drugs with delayed effects, such as those that act on kinases or DNA receptors, the buffering between drug concentration and effect is even greater. Less difference is seen in the peak-to-trough effect with multiple dosing, and the delay between peak drug concentrations and peak drug effect makes the drug more forgiving to poor compliance.

Time-Dependent Drug Effects

Homeostasis is one of the central properties of mammalian systems - the body responds to stimuli by trying to maintain equilibrium. Body temperature, blood volume, and the concentration of many ions in the blood are all examples of bodily functions that are tightly regulated. Changes in any of these will cause the body to respond to counteract the change. The same holds for drugs. The body will try to maintain homeostasis when a receptor is continually stimulated or inhibited.

For some drugs, a phenomenon known as tachyphylaxis occurs. Tachyphylaxis is the rapid development of tolerance to a drug after its administration. In practice, this means the drug's maximal effect decreases over time, which may sometimes be overcome by a higher dose. A classic example is the use of over-the-counter decongestants, such as pseudoephedrine or phenylephrine. Both these drugs stimulate α-receptors in the nasal cavity, pseudoephedrine indirectly and phenylephrine directly, producing local vasoconstriction of blood vessels. As such, these drugs are useful for treating runny noses that might occur with allergies or a cold. Continued use of these drugs leads to receptor internalization from the cell surface into the cell (down-regulation) and to decreased receptor responsiveness, both of which contribute to drug tolerance. Cessation of these drugs can then lead to rebound congestion because, as the body tries to counteract a runny nose through normal physiological mechanisms, there are few to no α-receptors available to act on. Hence, rebound congestion can occur until the body returns to its normal state.

Another example is the "high" seen with drugs of abuse after their administration. Figure 54 presents the "high" seen after intranasal cocaine administration in 4 volunteers as a function of cocaine plasma concentrations. The clockwise hysteresis loop provides evidence for tachyphylaxis. Look at cocaine concentrations at 15 min and 240 min after administration. They are about the same concentration. For a direct-acting, rapid-equilibrating drug, the effect at 15 minutes and 240 minutes should be the same if the concentrations are the same. That is not the case here. The "high" at 240 min is practically zero compared to 15 min after dosing. Volunteers don't experience any "high" at 240 min, even though they have the same plasma cocaine concentration as at 15 min after dosing.

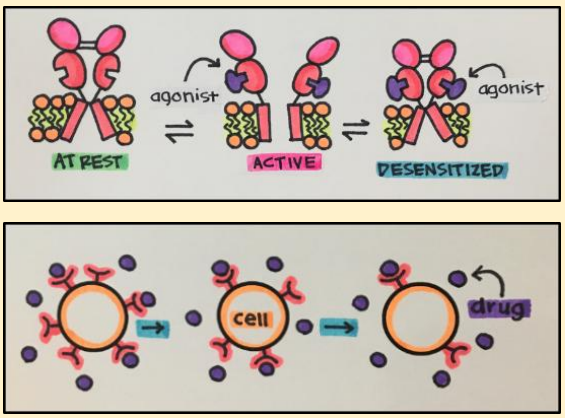
While short-term adaptation is called tachyphylaxis, prolonged use of some drugs, particularly drugs of abuse like alcohol, cocaine, and heroin, can lead to tolerance, which is a long-term decrease in effect over time. The difference between tachyphylaxis and tolerance is whether the decrease in effect after drug administration is dose-related. Tachyphylaxis does not appear to be dose-dependent, whereas tolerance

Tolerance can occur to both the desired effects of a drug and to its side-effects

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

is. Furthermore, tachyphylaxis has a more rapid onset than tolerance. An example of tolerance for a non-abused drug is the inhalational administration of β2-agonists, like albuterol or salbutamol, for the treatment of asthma. Asthma affects millions of people, and agonists mimic the effects of adrenaline in the lungs by stimulating 2-receptors, causing relaxation of central and peripheral airway smooth muscle and allowing dilation and facilitating oxygen exchange. Today, β-agonists are a mainstay of asthma therapy. Between 1959 and 1964, the incidence of asthma deaths in 5- to 34-year-olds in the UK increased 3-fold. This was attributed to a formulation of inhalational isoproterenol that was 5 times as potent as the traditional isoproterenol metered-dose inhaler (MDI). Further studies concluded that continued and prolonged use of β-agonists was responsible for the deaths.

Mechanisms of Tolerance



Two mechanisms of tolerance are desensitization and receptor down-regulation. With desensitization (top figure), the receptor becomes desensitized to the drug's effects to prevent overstimulation. In down-regulation (bottom figure), the cell internalizes its own receptors, thereby reducing the number of receptors available for the drug to bind to. Figures reprinted from Principles of Pharmacology-Study Guide, Copyright © by Edited by Dr. Esam El-Fakahany and Becky Merkey, MEd.

Pharmacological studies have shown that receptor downregulation and receptor desensitization are the mechanisms responsible. During an asthma attack, when bronchodilation is needed, the drugs have no effect because the receptors are not present to act on, and those receptors that are still available are less sensitive than drug-naïve receptors. Hence, the patient is unable to recover from the attack and may die as a result. Further studies suggest that a genetic component also plays a role. Today, long-term use of β-agonists carries strict contraindications for prolonged and continuous use.

The opposite of tolerance is drug sensitization. Seen with some drugs of abuse, evidence suggests it's a learned phenomenon and is what the word implies – drug effect increases with successive exposures to the drug. While it is difficult to demonstrate in humans, it can be readily shown in animals. Rats, when dosed with cocaine, show increased locomotor movement. Rats intermittently exposed to cocaine in the same location will show increased locomotor activity with subsequent doses. Yet at the same time, rats continuously dosed with cocaine will show tolerance. The location component of the regimen is central to sensitization, as rats previously dosed at one location and then at another fail to show sensitization. Hence, sometimes this phenomenon is referred to as 'behavioral sensitization.'

Related to behavioral sensitization is the context-dependent effect of drugs. Sensitization is only observed in the environment where the drug has been previously administered. For example, only animals dosed in their home cage may demonstrate sensitization; animals previously dosed in their home cage when given the same dose in a novel environment do not. Context-dependent pharmacodynamics is more general in that the effect of a drug can change depending on the context in which it is given. This effect is seen with many drugs of abuse. An example that everyone might be familiar with is the effect of alcohol. Alcohol is a depressant; it shows cross-tolerance with benzodiazepines, but in the right context, like at a party, alcohol acts as a stimulant by releasing behavioral inhibitions, and drinking alcohol while in a bad mood often leads to making that mood even worse. Hence, context can play a role in the effect of some drugs.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drugs that show tolerance, tachyphylaxis, or sensitization will also show cross-tolerance and cross-sensitization. None of these are drug-specific. Drugs that activate the same receptor, either directly or indirectly, will show cross-tolerance or cross-sensitization to the responsible drug. For example, people who are tolerant to cocaine will show cross-tolerance with methamphetamine, or people who are chronic alcohol abusers will show cross-tolerance to benzodiazepines. Likewise, sensitization to cocaine will show sensitization to amphetamine.

As with downregulation in the case of continuous agonist stimulation, receptors may also undergo upregulation in response to continuous antagonism. Although the exact mechanism is unknown, it is believed that tardive dyskinesia, seen with prolonged use of antipsychotics (which are dopamine antagonists) used in the treatment of schizophrenia, is due to up-regulation of dopamine receptors. Chronic blockade of dopamine receptors leads to upregulation, which may cause potentially permanent involuntary movements, such as lip smacking, tongue thrusts, and eye blinking. This theory is not entirely complete because tardive dyskinesia can continue after discontinuation of antipsychotic use.

A particular area where tolerance can be life-threatening is cancer. Many cancers have developed sophisticated mechanisms that overcome the effects of drugs and prevent them from working, a process known as drug resistance. Patients may see their tumors shrink or completely go away, but then, with continued treatment, their cancer will return, their tumors will increase in size, and this time, the tumors will be unresponsive to further treatment by the drug. Altering the metabolism of the drug, changing the levels of the drug's target, compartmentalizing the drug so it can't reach the target, decreasing the rate of a drug's uptake into the tumor cell, and increasing the rate of drug efflux from the tumor are some of the mechanisms cancer cells use to increase their resistance to cancer drugs [152]. When a tumor becomes resistant to a drug, a new drug with a different mechanism of action is often required to treat the cancer. With each successive failure, it becomes increasingly more difficult to treat the cancer.

Chronopharmacokinetics and Chronopharmacodynamics

Mammals possess an internal clock, often synchronized with the light-dark cycle, that regulates numerous biological processes. Such a clock creates a circadian rhythm in which the activity and concentrations of many macromolecules, such as enzymes and receptors, change during the day. For example, in most people, cortisol (the body's natural stress hormone) concentrations rise in the morning, peak around 9 am, and then decrease throughout the day, reaching a nadir around midnight. The cycle repeats itself the next day. Just as many biological processes are affected by circadian rhythms, so are disease processes. For example, a runny, sneezy nose is often worse in the morning after a night's sleep; the incidence of heart attack is greatest in the morning, whereas congestive heart failure is typically worse at night. It may come as no surprise that circadian rhythms influence both the pharmacokinetics and pharmacodynamics of certain drugs, and that for some drugs, dosing is recommended at a specific time of day.

Take statin dosing, for example. Cholesterol is an important component of cell membranes and serves as a precursor for steroid hormones and bile acids. But if cholesterol concentrations are too high, the risk of cardiac events like heart attacks and stroke increases through the development of atherosclerosis. HMG-CoA reductase is the rate-controlling enzyme in the synthetic pathway that creates cholesterol in the body. Statins are a group of drugs that reversibly inhibit HMG-CoA

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

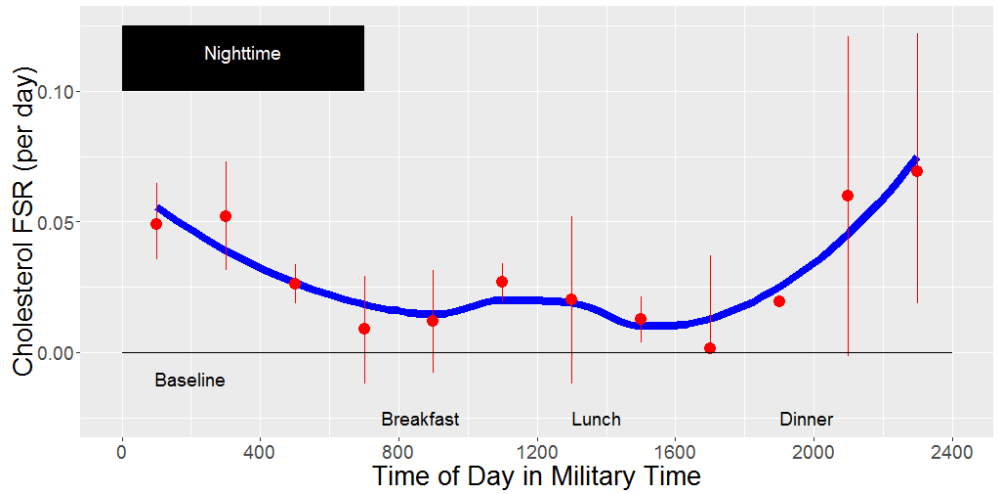


Figure 55. Fractional cholesterol synthesis rate (FSR) versus time of day. Cholesterol synthesis follows a diurnal rhythm, with peak synthesis at night. FSR is the cholesterol synthesis rate relative to the total pool of cholesterol in the body. Figure modified from [153].

reductase, thereby decreasing blood cholesterol concentrations and reducing the risk of cardiovascular disease and mortality. The major statins are Mevacor (lovastatin), Zocor (simvastatin), Lipitor (atorvastatin), Pravachol (pravastatin), and Crestor (rosuvastatin). Both simvastatin and lovastatin recommend that their drugs be taken at night. The package insert for simvastatin states, "Recommended usual starting dose is 10 or 20 mg once a day *in the evening*", while lovastatin's label states, "The usual recommended starting dose is 20 mg once a day given *with the evening meal*."

Why do both products recommend nighttime dosing? It turns out that a circadian rhythm regulates cholesterol synthesis. Cholesterol synthesis is at its peak in the nighttime hours and at its lowest in the early morning and mid-afternoon (Figure 55). To have the greatest effect on cholesterol synthesis, the best time of day is to inhibit HMG-CoA reductase activity at night. But then, why do the labels for atorvastatin, pravastatin, and rosuvastatin not mention nighttime dosing? Atorvastatin's label reads, "Recommended start dose: 10 or 20 mg once daily," whereas pravastatin's label reads, "The recommended starting dose is 40 mg once daily". Why are the labels different? The answer lies with pharmacokinetics. If you compare the elimination half-life of each product and its dosing recommendations, a clear distinction is made:

Statin	Half-life (h)	Time of Day Dosing
Lovastatin (Mevacor)	1-2	Nighttime
Simvastatin (Zocor)	3	Nighttime
Atorvastatin (Lipitor)	14	Anytime
Pravastatin (Pravachol)	77	Anytime
Rosuvastatin (Crestor)	19	Anytime

Those drugs that have nighttime dosing have very short half-lives. If the drug is taken in the morning or at night, it will be cleared from the body and won't be able to affect cholesterol synthesis when it is at its nighttime peak. Hence, dosing of these drugs must be at night to have the greatest effect on

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

HMG-CoA reductase. On the other hand, drugs with dosing at any time of day have longer half-lives and maintain sufficient concentrations in the body to affect cholesterol synthesis when it is at its nighttime peak. Atorvastatin also has active metabolites with half-lives ranging from 20 to 30 hours, which contribute to its inhibitory effect and further support the use of anytime dosing.

And Now For Something Completely Different

- **Chronopharmacokinetics** studies the effect of circadian rhythms on the pharmacokinetics of drugs.
- **Chronopharmacology** studies the effect of circadian rhythms on the pharmacodynamics of drugs.
- **Chronotherapeutics** is the alignment of drug timing to optimize effect based on time of day.

Earlier in the chapter, it was mentioned that the mechanism of action for most drugs is mediated through protein-drug interactions, such as those involving receptors or enzymes; however, many drugs do not follow this pattern. While it would be impossible to cover every single one of these, some are worth mentioning. These will be highlighted in this section.

First are the drugs that work by chemical or physical mechanisms. Osmotic diuretics, such as mannitol, act nonspecifically by increasing plasma osmolality; in response, the kidney increases urine output to maintain homeostasis. These drugs work by physical mechanisms. Antacids increase stomach pH nonspecifically, and chelating agents are used to treat accidental poisoning from orally ingested toxins by chemically binding the toxins. These drugs work by purely chemical mechanisms.

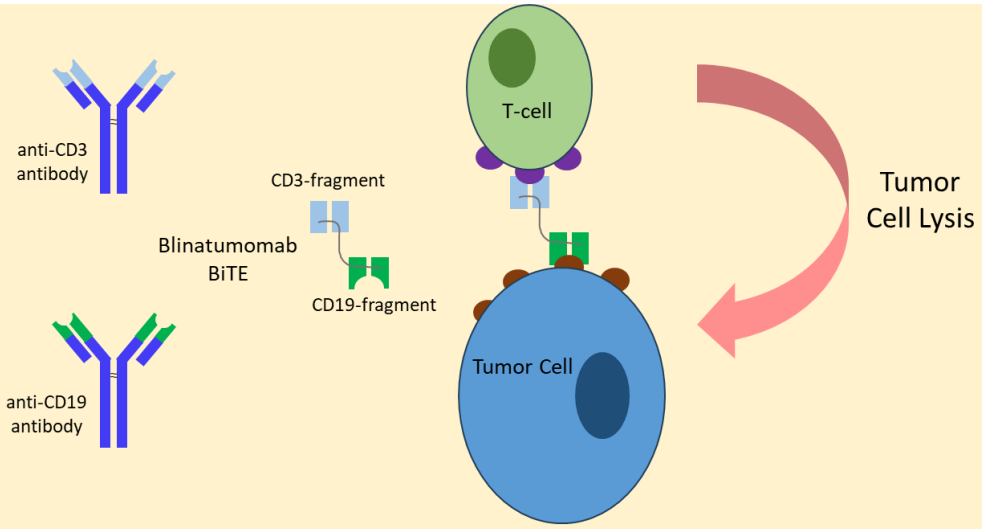
Biochemists today are becoming increasingly adept at combining products. Antibody-drug conjugates, used to treat various cancers, combine a monoclonal antibody with a small molecule chemotherapy drug, known as the payload. The antibody binds to its receptor on the tumor cell, and once inside the cell, the small molecule is released, leading to tumor cell death. The specificity of the monoclonal antibody reduces the side-effect profile typically associated with the chemotherapy payload. Today, ADCs are being developed not only with small-molecule chemotherapy drugs but also by combining with topoisomerase inhibitors, fungal exotoxins, tubulin binders, and photosensitizers [154].

Bispecific antibodies are derived from monoclonal antibodies, but instead of having both antibody arms bind to the same antigen, they have different antigen-binding arms that bind to two distinct antigens. Unlike monoclonal antibodies, which induce cell death through various mechanisms, bispecific antibodies are not designed to kill tumor cells directly. They are designed to bring things together. For example, Blincyto (blinatumomab) is an immunotherapy that has two binding arms, one for CD3 and the other for CD19, both of which are proteins located on the cell membrane. CD3 is on T-cells and CD19 is on leukemia cells. When blinatumomab binds to CD3 on T cells and CD19 on leukemia cells, the leukemia cell and T cell are brought into close proximity, thereby facilitating the T cell's ability to kill the tumor cell. There are dozens of other bispecific antibodies and derivatives currently in development. As of 2024, 12 have been approved by the FDA.

Oligonucleotides don’t work on proteins, like most other drugs; instead, they are small nucleic acid sequences (usually 6 to 30 in length) that regulate gene expression. These include:

- antisense oligonucleotides that block the function of mRNA and interfere with protein production,
- small interfering RNAs (siRNAs) that degrade specific mRNA and reduce protein production,
- microRNAs (miRNAs), which bind to mRNA and prevent protein translation from occurring.
- mRNA vaccines, such as the COVID-19 vaccine, that act as endogenous mRNA and code for the production of specific proteins, and
- aptamers, which are DNA or RNA molecules that bind to specific targets, such as proteins or small molecules.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Is blinatumomab an antagonist? We think of monoclonal antibodies as antagonists of a target, blocking the action of an antigen. Bispecific antibodies, such as blinatumomab, are antibody fragments that contain only the binding domain of a monoclonal antibody and bring a T-cell into proximity with a tumor cell, allowing the T-cell to kill the tumor cell. Bispecific antibodies that engage T cells are called Bispecific T-cell Engagers (BiTEs). They are not antagonists but rather matchmakers, bringing together the T cell and the tumor cell. There are many approved BiTEs on the market now, and what distinguishes them is the tumor cell target. Tecvayli (teclistamab) and Elrexfio (elranatamab-bcmm) both use CD3 and BCMA, B-cell maturation antigen found on multiple myeloma cells. Columvi (glofitamab-gxbm) uses CD3 and CD20, which are found on B-cell lymphoma cells.

The market for oligonucleotides is projected to reach over $30 billion by 2030.

Lastly, there is a plethora of exogenous, naturally occurring protein supplements approved for use [155]. These include insulin, erythropoietin, clotting factors, gonadotropins, growth hormone, thyroid hormone, and recombinant enzymes. These proteins are designed to supplement or boost protein levels when they are deficient or low. Due to their poor stability, some of these compounds may be chemically modified, thereby increasing their stability as a formulation or within the body. They work by binding to cell surface receptors and triggering intracellular effects, similar to their natural counterparts. Most of these products are administered non-oral by injection, and significant research is being invested in finding ways to allow oral administration.

Summary

Pharmacodynamics is a complex subject that has been studied for more than 100 years, with the first reported observation of a "receptive substance". Even today, receptor theory does not completely explain everything, so the field continues to evolve. These last two chapters introduced the basic concepts of pharmacokinetics and pharmacodynamics, the backbone of clinical pharmacology. Upcoming chapters will discuss how a drug's pharmacokinetics and pharmacodynamics vary across patients, and how a dosing regimen may need to be adjusted accordingly.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026.

Chapter 6

Drug Development

This chapter will introduce the drug development process, and although it is presented as a linear, fixed process, it is not. The process depends on the drug, the proposed indication, whether it can be administered to healthy volunteers or only to patients, and the degree of risk the company is willing to assume. For instance, the company may decide not to do certain studies that the FDA may later require. Most approved drugs are small molecules. In 2024, 2 out of 3 FDA-approved drugs were small molecules. As such, this chapter will focus on the development of small molecules (i.e., chemically synthesized drugs) in healthy volunteers, and the differences between biologics development and the development of a patient pathway drug will be noted. Cerdelga® (eliglustat), an orally administered drug used to treat Pompe disease, a rare disease that weakens muscles throughout the body, will be used as a prototypical drug to illustrate the process.

The Process of Drug Discovery to Lead Optimization

Early drug discovery was chemocentric, built around individually synthesized chemicals. Pharmaceutical companies employed hundreds of chemists to synthesize potential new drugs. The focus was on the product's chemical nature. I am calling it a chemical at this point because there is no evidence that it will have a pharmacologic effect and be considered a drug. Once the new chemical was synthesized, it was screened in various animal models and receptor assays to assess its activity. This process, called a phenotypic screen, is target-agnostic; it casts a wide net to see if the drug has activity at any receptor. If activity was discovered, the chemical moved through the pipeline for further development. This process was inefficient and could not be sustained over time. Around the 1960s, modifications were made to the process to enhance its efficiency, such as conducting focused screening or focused synthesis around known chemical structures (called pharmacophores) that exhibited activity in specific indications. But, by and large, the process remained inefficient.

With the development of modern biology, drug discovery switched in the 1990s to a more focused, targeted approach (Figure 56). Researchers would target drugs to proteins or other macromolecules associated with a disease, the idea being that drug discovery would be more "rational" and more efficient, a process that appears to be working. Although an analysis of all approved drugs from 1999 to 2008 suggested that more drugs were approved using a chemocentric approach, a more recent analysis of 113 first-in-class drugs approved by the FDA reported that from 1999 to 2013, 78 of them were discovered using a target-based approach [156]. If you consider that all biological drugs are target-based and remove these from the analysis, 45 of 80 small molecules were discovered using a target-based approach.

Today, there is debate over whether targeted approaches are improving the efficiency of drug discovery. Sadri [157] analyzed 32,000 published articles and patents going back 150 years to determine how individual drugs were discovered. He concluded that despite several decades of using a targeted approach to drug discovery, only about 9% of drugs were discovered using a targeted approach. He argues that companies should return to more phenotypic screening to identify potential

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026.

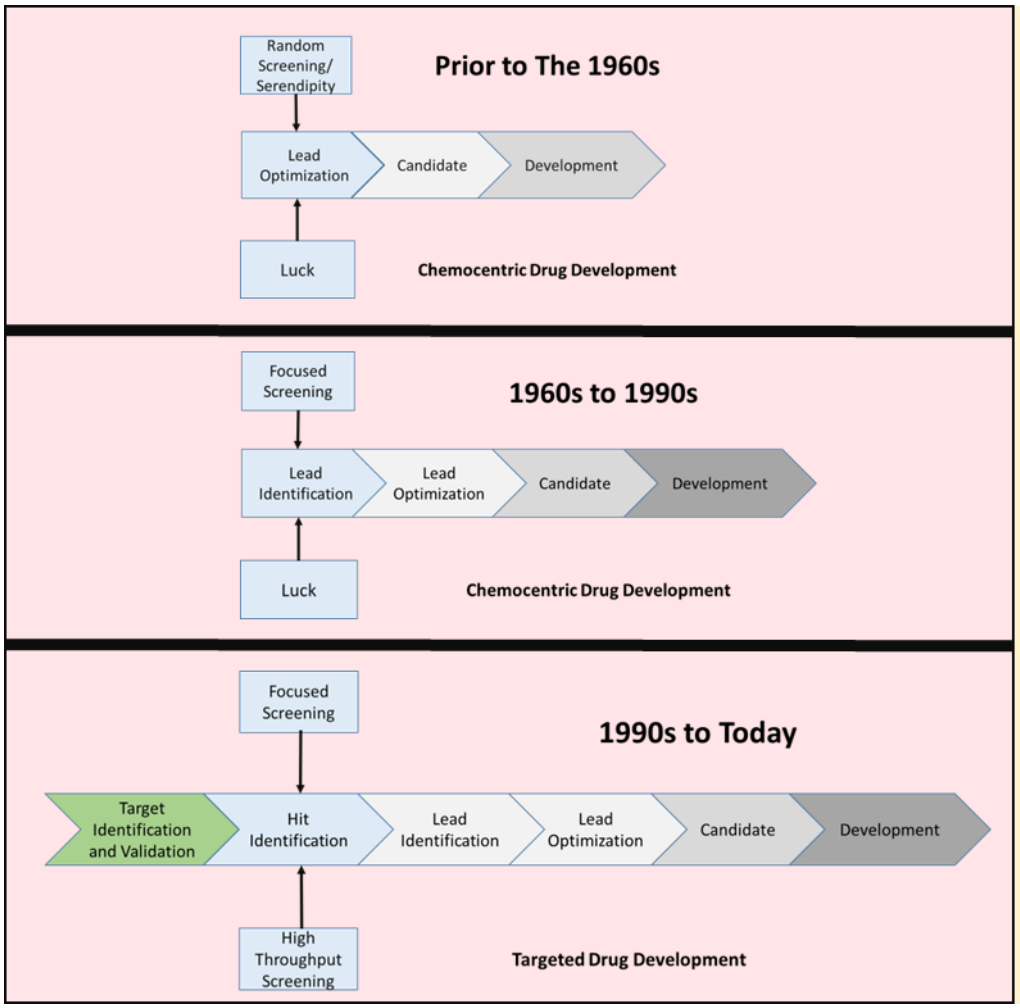


Figure 56. Drug discovery paradigms by era. Drug discovery before the 1990s was chemocentric in that drugs were discovered, and indications were found to fit the drug. In the 1990s, that all changed and became target-centered. Once an indication and a target were identified and validated, compound libraries were screened, or focused screening was used, to identify an appropriate lead, which was then optimized and tested in humans for study.

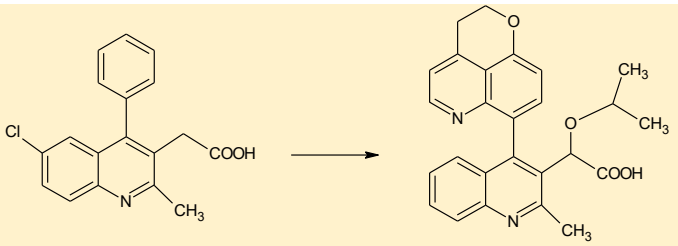
drug candidates better.

Nevertheless, today most drug companies take a targeted approach to drug discovery, beginning with identifying and validating targets. Starting from a disease of interest, researchers will examine changes in the pathophysiology of patients with the disease compared with healthy subjects, seeking alterations in proteins, transporters, receptors, and other relevant molecules. This is target identification. Once a change is identified, such as a change in protein concentration, researchers work to confirm that the change is specific to the disease and not a nonspecific effect, and that modulating the target is likely to have therapeutic benefits. This is target validation.

For example, researchers may identify that a specific protein is elevated in a disease. Then, using genetic methods, they may create an animal model, such as a mouse, that lacks the gene encoding the protein. Does the disease process still occur in knockout animals? If so, that would suggest the target is not disease-specific. The problem with target identification and validation is that most

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

diseases are complex, multifactorial, and not the outcome of a single process. Hence, each target does not act with equal weight in the disease process, and not all are equally effective. It was once believed that if a target was validated, then creating very specific drugs for that target may be more beneficial and have a better safety profile than less specific drugs. But this is not necessarily the case.

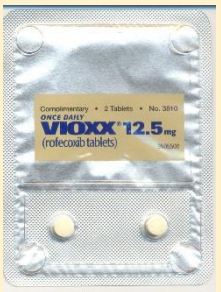


HTS Hit #1
EC50 > 40 uM

BI 224436
EC50 11-27 nM
Potency 2,000x more selective

Example of a Hit to Lead for a new HIV Integrase Inhibitor. Hit #1 was identified by high-throughput screening (HTS). Subsequent testing and chemical modification led to the discovery of BI 224436, which was shown to be 2,000-times more potent than HTS Hit #1 [158].

Consider the COX-2 inhibitors. Cyclooxygenase (COX) is an enzyme involved in the synthesis of prostaglandins, which are known mediators of pain. COX exists in two forms, Type I and II. COX-1 mediates prostaglandin synthesis in the stomach and protects the stomach from ulcers. COX-2 synthesizes the prostaglandins responsible for pain. Aspirin inhibits both COX-1 and COX-2, and, as expected, has anti-inflammatory effects; however, its long-term use may lead to the development of stomach ulcers. By creating specific COX-2 inhibitors, it was believed that patients would experience pain relief without developing stomach ulcers. Vioxx (rofecoxib) was a specific COX-2 inhibitor compared to aspirin; however, it was later shown to increase cardiovascular risk and was withdrawn from the market.



Vioxx was a specific COX-2 inhibitor thought to have anti-inflammatory effects without any gastrointestinal effects. However, Vioxx was shown to increase cardiovascular risk, leading to heart attacks and strokes. It was withdrawn from the market in 2004 due to safety concerns. Subsequent studies have shown that COX-2 is present in the cells lining blood vessels, and that blocking it predisposes mice to blood clotting and hypertension. Additionally, when prostaglandin synthesis is blocked, precursors used to produce prostaglandins are diverted into another biochemical pathway, leading to the formation of leukotrienes, which have been linked to atherosclerosis.

Validating a target requires effort and may take decades. It requires confirmation from multiple sources, and even then, questions may still arise. The problem stems from the fact that the biology behind many diseases remains largely misunderstood. Alzheimer's, diabetes, and schizophrenia are all examples of common diseases that scientists still don't understand the causes of or how to prevent. Therefore, targets are typically classified into two groups: putative and validated (or established). And even if a target is validated, it may not be 'druggable' in the sense that it is available to a drug. In other words, a drug may not be able to bind to the target with any degree of specificity.

Once a target has been identified, compound screening may begin. Thousands of compounds from different chemical libraries may be screened to identify those that bind to the target with high affinity. This step is called lead generation, also known as 'hit-to-lead'. Those compounds that meet the affinity threshold provide a basis for further

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Use of Genetics to Improve Target Validation

The influence our genes have on our development varies. In some cases, genes are destiny. Some diseases are autosomal, meaning they are passed down from one or both parents on non-sex chromosomes. Autosomal dominant or recessive traits tend to have a strong genetic influence. Most diseases, however, result from the interaction of genetics and environmental influences. Based on data from identical twins, diseases like leukemia have a low disease risk attributable to genetics (3%). In contrast, others, such as asthma (49%), were high, leading some to conclude that genetics is not the major cause of chronic diseases [159]. Nevertheless, it has been estimated that the probability of successfully launching a drug is 2.6-times more likely if the drug's proposed mechanism of action has genetic support behind it. However, the authors do acknowledge that such evidence is often retrospective, not discovered until late in the development program or after drug approval, and was often not the impetus for development of the drug in the first place [160].

Because the contribution of genetics to disease is often low, scientists use genome-wide association studies (GWAS) to identify the links between genes and disease. GWAS involves screening thousands of patients, each with thousands to millions of DNA nucleotides in their genetic sequence, and then using complex numerical methods to identify which nucleotides and genes might be important. By comparing the genetic markers of those with the disease to those without, GWAS has been used to identify the genetic basis of many diseases, such as Parkinson's, obesity, and cardiovascular disease. The difficulties with GWAS include the high likelihood of falsely identifying genes as disease-associated and the fact that correlation does not necessarily imply causation. A gene being associated with a disease does not mean that the gene causes the disease. Furthermore, genetics alone does not account for environmental influences, which can be mediated through epigenetic modification. An example of this is when DNA becomes methylated, silencing the gene it encodes.

Converting evidence of a genetic association into a drug is challenging [161] and how companies use genetic changes as a basis for target validation depends on the company. Some companies use it early in their target nomination process. An example of this is Regeneron's REGN5381, a monoclonal antibody agonist being developed for the treatment of heart failure. Starting with a GWAS of heart failure in 700,000 patients, Regeneron identified variants in the NPR1 gene and developed an antibody that alters the conformation of the NPR1 receptor, thereby improving hemodynamic efficiency [162]. Others, however, have said that relying on genetics as a threshold for candidate nomination may miss opportunities for drugs that might otherwise have had a high probability of success [161]. For this reason, many companies use the totality of evidence to validate a drug target, rather than relying solely on one modality.

optimization.

Computer models and experiments, employing trial and error, can be used to identify the parts of the molecule responsible for binding to the target. This is referred to as the pharmacophore of the molecule. Once the pharmacophore of the molecule is identified, it can be chemically modified to create better molecules with improved potency against the target or to enhance their pharmacokinetic properties.

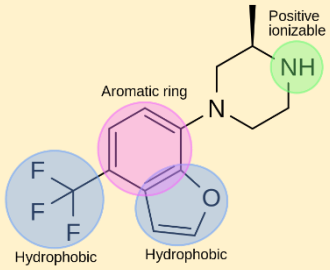
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Further, studies in biochemical or animal systems may be used to examine the biological activity of the lead. If the lead is unsuitable, chemists may further refine the lead until a suitable drug candidate is found. This iterative process is called 'lead optimization.' ***Not every compound that is synthesized can be a suitable drug candidate.*** One problem with high-throughput screening (HTS) is that it often favors lipophilic compounds. But, too high of lipophilicity and the drug will not be a suitable candidate for oral administration. Christopher Lipinski and colleagues at Pfizer examined a database of ~2,500 small molecule drugs and identified properties consistent with "a drug" [114].

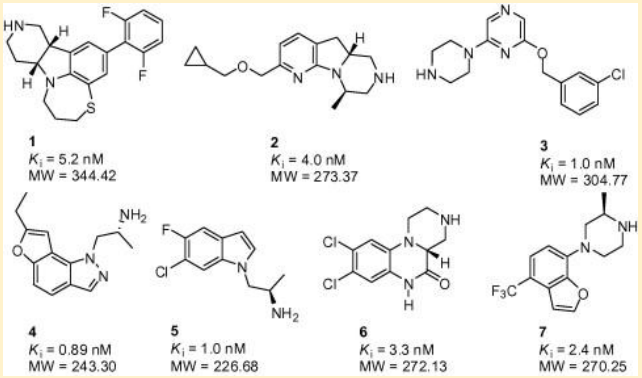
These properties became known as "the rule of 5" (it's not the number of rules, but that the number '5' is in every rule):

1. Molecular weight < 500 Daltons;
2. The total number of hydrogen-oxygen and hydrogen-nitrogen bonds less than 5;
3. The total number of nitrogen and oxygen atoms is less than 10; and
4. Octanol-water partition coefficient < 5.

All the rules are multiples of five – hence the name. Compounds that had rule violations were unlikely to have the properties necessary to become "a drug." For years, companies screened compound libraries to ensure no rule violations before taking a compound to the candidate stage. Since its release, there have been many variants of the rule and alternative sets of rules, there have been debates about the cutoff values (e.g., whether the cutoff should be 400 Daltons instead of 500 Daltons), and whether even the rule of 5 still applies to the drugs being developed today, but for the most part, Lipinski's rule of 5 still provides a quick rule of thumb to determine whether a compound could be an orally administered drug. A 2001 study found that 90% of orally administered drugs had no more than one rule violation of the Rule of 5. Later estimates put this number closer to 80 to 90%.



Pharmacophore model for 5HT2-C receptor agonists. Structure-activity relationships have demonstrated that specific binding to the 5-HT2C receptor is achieved when a planar hydrophobic region is attached to a hydrophobic side chain, and when an ionizable side chain is positively charged. Below are seven proposed 5-HT2-C agonists developed by Asif Ahmed and colleagues that retain the pharmacophore [163].

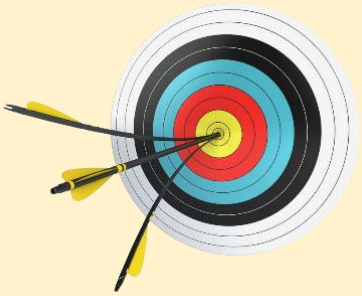


Hits have a proposed link between their molecular structure and the target's activity.

A **Lead** is a compound series derived from a **Hit** that has a link between its structure and activity in cell-based or biochemical systems that may be therapeutically beneficial. Leads may require further optimization.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Target Product Profile



The Target Product Profile (TPP) is a strategic tool used by pharmaceutical companies to define the minimum acceptable characteristics and performance features a drug must have for commercial success. The TPP is defined early in the drug development process and serves as the guiding light for the development plan. For example, suppose a competitor drug is taken once a day, then one criterion on the TPP for your new drug may be 'Once-Daily Oral Dosing.' If it is shown that the half-life of your drug is short and has to be taken twice daily, this is a killer because it is unlikely you will be competitive in the marketplace, and the drug may be terminated from further development as a result. Often, the TPP is built around the desired label and includes the intended use, target population, safety/efficacy requirements, pharmacokinetic characteristics, and pharmaceutical properties. Recent developments to the TPP include patient input into the profile.

As an example of a TPP, consider a hypothetical diabetes drug. Its TPP might be:

Property	**Target**	**Minimum**
Indication	Adjunct to diet/exercise in adults	Same
Efficacy	6-month Hgb A1c reduction of 1.5% or more vs. placebo	6-month Hgb A1c reduction of 1% or more vs. placebo
Safety	No hypoglycemia risk	<5% incidence of hypoglycemia; no Grade 4 hypoglycemia
Dosing	Once daily oral tablet	Once daily oral tablet
Differentiation	Weight decrease between 10% to 20%	No change in weight
Regulatory Goal	No MACE study needed	Standard Phase 3 study with MACE endpoints

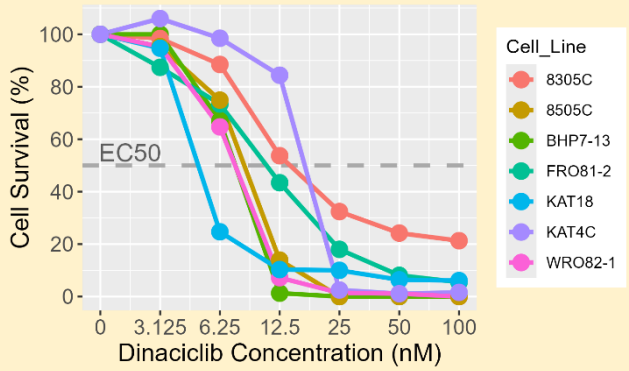
The TPP is designed to be a living document and may be updated along the way. This is one of the problems with the TPP concept – failure creep. Returning to the diabetes drug example, suppose that at the end of the study, the drug decreased Hgb A1c by 0.9%, which is below our minimum criteria for success. You've just spent millions of dollars on clinical trials and spent years getting to this point, do you just throw it all away and terminate the drug from further development? Many companies may push forward regardless of the TPP, which may explain the high failure rate in Phase 3 due to a lack of efficacy.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Preclinical Development

Once a compound progresses to the candidate stage, it then enters preclinical development, which consists of four activities:

1. Preclinical in vitro and in vivo pharmacology,
2. Preclinical pharmacokinetics,
3. Preclinical toxicology and safety, and
4. Formulation development.



Example of a Typical Dose-Response Profile. Dinaciclib induces cytotoxicity in thyroid cancer cells. Shown here is the percentage of cells surviving after 4 days of incubation with different concentrations of dinaciclib in different cell lines. These are typical dose-response profiles – as drug concentrations increase, the effect of the drug increases to a plateau (in this case, the effect decreases to a nadir, with no cells surviving). Figure redrawn from [164].

The first preclinical studies done are in vitro and *in vivo* pharmacology. Pharmacologists will dose animals with the drug or use in vitro models like cell culture to determine if there is a drug effect. They will use different concentrations or doses of the drug to determine the dose-response profile. They will also compare the new drug's efficacy to that of potential competitors to determine whether it is more potent or more effective. A key feature of these studies is to discern or confirm the mechanism of action. If a drug is thought to affect a particular receptor but does not bind to it, then how the drug works may need to be rethought and confirmed through experimentation.

Other scientists may begin to determine what is the fate of the new compound when it is administered to animals. How is the drug absorbed? Once it is absorbed, where in the body does it distribute? Does the compound concentrate in any particular tissues? When the drug is cleared from the blood, do tissues still contain it, indicating irreversible binding to tissue proteins? How is it metabolized? If you compare the metabolism using isolated liver cells from rats, mice, dogs, and humans, are there differences in the drug's metabolism? Do certain species make metabolites that other species do not? How is the drug and its metabolites eliminated from the body? Does the liver predominantly metabolize them, or are they excreted in the urine? If you change the formulation, how are the drug's pharmacokinetics affected? These are the kinds of questions a preclinical pharmacokineticist would answer.

A key set of studies conducted before testing in humans are toxicology studies, which are designed to assess the safety of a compound. Toxicology studies give high doses of the drug to maximize the possibility of observing some adverse event. Many different types of toxicology studies are done during the drug development process:

- Single and multiple doses of the drug using the intended route of administration in rodents and non-rodents;
- Genotoxicity studies to determine whether the compound causes DNA damage;
- Carcinogenicity studies to determine whether a compound causes cancer;
- Reproductive toxicology studies to determine whether a compound affects a fetus in utero;
- Safety toxicology studies to determine whether the drug affects heart function, laboratory tests, or animal behavior.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Depending on the drug and the intended indication, other specialized toxicology studies may be needed:

- Local tolerability studies. If a drug were to be given cutaneously via the skin, rabbits might be used to see if the drug and vehicle (patch, cream, etc.) cause irritability of the skin;
- Juvenile toxicology studies. Before a drug is studied in children, juvenile animals will be administered the drug in an attempt to predict toxicity in children.

"Two critical "wrong" decisions regarding animal tests of human pharmaceuticals are:

1.) *falsely identifying a toxic drug as 'safe' and*

2.) *falsely labeling a potentially useful therapeutic agent as 'toxic'."* –

Gail A. van Norman [165]

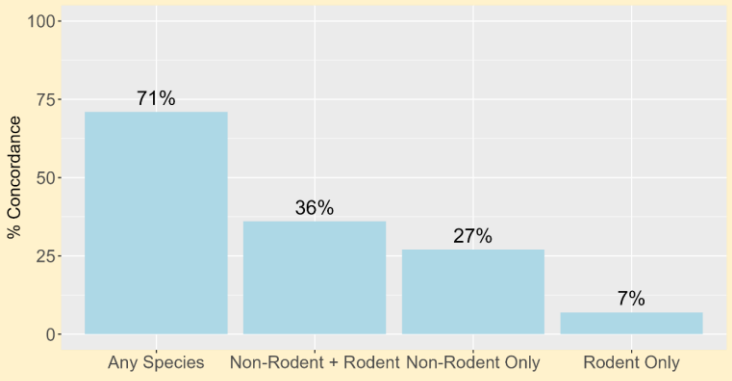
Selection of the best animal model predictive of human toxicity is critical because it is possible some animals might not have the target (an example is testing compounds for prostate cancer in female rats or if the disease is very different in males compared to females), the target may not react the same in animals as humans (an example is that unleaded gasoline produces a unique kidney injury in rats that is not seen in humans because of accumulation of a protein only found in rats), or that there might be a difference in how drugs are metabolized between humans and animals (an example is tamoxifen, which is genotoxic in rats and mice, but not humans, because rodents make a unique metabolite not seen to any significant extent in humans).

Consequently, when interpreting the results of toxicology studies, toxicologists may ask some of the following questions to understand the relevance to humans:

- Was the appropriate animal model used?
- Were the dose(s) and dosing duration sufficient to support the proposed clinical study or labeling?
- Were adequate systemic blood concentrations achieved in animals?
- Was the route of administration relevant to clinical use?
- Did the test system exhibit any effects?
- Were the effects treatment-related? Dose-related?
- Were the effects biologically significant?
- Were the effects reversible?
- Were the effects clinically relevant?
- Can the effects be monitored clinically?

Translation of animal results to humans is challenging, but toxicologists are making significant progress in enhancing the reliability of their models.

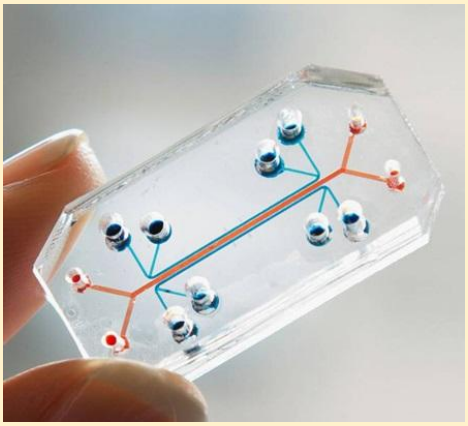
The downside of toxicology studies is that they are not perfectly predictive of toxicity in humans. For example, an animal may show cardiovascular



Concordance between the toxicity seen in animals and humans. Rodents were mice, rats, guinea pigs, and rabbits. Non-rodents were dogs, primates, or both. Concordance was declared if similar toxicity was observed between humans and animals in the same organ system, e.g., gastrointestinal or cardiovascular [166].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

toxicity, but the same may not occur in humans. The best overall concordance between animals and humans is 71% [166]. This value may vary depending on the organ system class. The highest concordance is seen with gastrointestinal, hematologic, and cardiovascular toxicity, and the lowest is seen with cutaneous and hypersensitivity reactions. These differences may be attributed to various factors, including differences in drug metabolism or variations in the target structure between animals and humans, e.g., the structure of even common proteins, like albumin, differs between species.



Microphysiological systems (MPS) are advanced in vitro models that utilize 3-dimensional cell cultures to mimic the functional structure of an organ, and employ microfluidic devices to simulate blood flow and nutrient delivery to the system. Examples of MPS systems are so-called "organ on a chip" systems, organoids, and multi-organ systems. Currently, there are many organ-on-a-chip systems, including brain, gut, lung, heart, kidney, and liver. Figure reprinted from Wikipedia Creative Commons.

No one likes the idea of using animals as guinea pigs to test new drugs. As such, tremendous efforts are being made to move away from animal testing. Since 1959, the idea of the "3Rs" in animal research has been promoted [167]:

- Replacement, such as using human cells or tissues to replace animal testing;
- Reduction, using a small number of animals in animal studies, and
- Refinement, using laboratory procedures that reduce the pain and distress to animals during animal studies.

In December 2022, the FDA Modernization Act 2.0 was signed into law [168]. It's a short bill that modifies the FFDCA of 1938, allowing the use of nonclinical models as an alternative to animal testing instead of requiring preclinical testing in animals before administration in humans. These nonclinical models are defined as:

> *'nonclinical test' means a test conducted in vitro, in silico, or in chemico, or a non-human in vivo test that occurs before or during the clinical trial phase of the investigation of the safety and effectiveness of a drug, and may include animal tests, or non-animal or human biology-based test methods, such as cell-based assays, microphysiological systems, or bioprinted or computer models.*

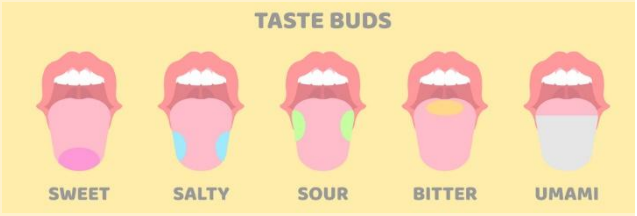
The bill also removes the requirement for animal testing in biosimilar development and allows alternatives to be used for assessing toxicity. In 2025, the FDA took a further step and introduced a roadmap to reduce animal testing using validated New Approach Methodologies (NAMs), such as computer simulation, machine learning, organoid cell cultures, and MPS systems [169]. This program will start with monoclonal antibodies but will expand to other drug classes in the future. In late 2025, the first case of using human vascularized organoid-based NAM combination studies, without relying on animal efficacy testing, to support the IND submission of BAL0891, a new checkpoint inhibitor, was reported [170]. While these efforts are encouraging and a good first step, we are far from entirely removing animal testing from the development process, and further work is needed before these alternatives are routinely used.

While all this is happening, formulation scientists are working with the compound to develop a suitable formulation for human administration. Formulation science converts an active pharmaceutical ingredient (API), also called the Drug Substance, into a quality Drug Product in a specific dosage form. The choice of formulation is driven by commercial necessity, patient convenience, and stability needs. For an oral formulation, is a capsule sufficient, or is a more stable tablet form needed? For example, in the case of a stroke victim who may be unable to move or

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Why Do Drugs Taste So Bad?

Have you ever accidentally broken a pill in your mouth before swallowing it? Or failed to swallow a pill and have it linger in your mouth too long before swallowing? Or perhaps you had to take a pill sublingually or buccally? In these instances, you know how bitter a pill tastes. But why do they taste bad?



Taste is one of the five senses that allow us to perceive the outside world. On our tongues are taste buds, a collection of 50 to 100 cells. Taste receptors are located within the taste buds and on the cell surface. These receptors enable us to perceive the flavors of sour, sweet, salty, umami, and bitter. The flavors are clustered on different parts of the tongue, such as sweet being located at the tip.

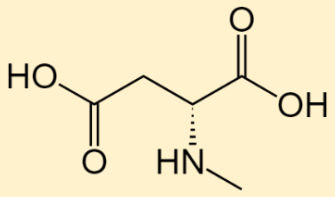
Because early humans ate plants, and many of these plants were toxic, the human body developed 27 different bitter receptors to help us identify when not to eat something [171]. This contrasts with the few distinct receptor types for sweetness. Some of these receptors undergo small structural changes due to genetic polymorphisms (more on this topic in the chapter on Pharmacogenetics and Pharmacodynamics). Because of these structural changes, they perceive bitter flavors as even more bitter than the rest of the population [172]. There is also some evidence that different ethnicities perceive bitterness differently. For example, when propylthiouracil was placed on the tongues, Asians perceived it much more bitterly than Europeans, Africans, or South Asians [112].

Many drugs administered orally are derived from plants and chemically activate these bitter receptors. This is why they taste bad. There is little that can mask the taste of a bitter drug. Encapsulation, adding other chemicals to mask the bitter taste, and using sweeteners have all been tried to varying degrees of success. But truly bitter drugs are hard to mask.

swallow, an oral formulation may not be a viable choice; instead, a rectal suppository or intravenous formulation may be better, but even this may not be easy to administer in the field.

Over time, a formulation may degrade, changing the strength of the drug in the formulation or may form unsafe degradation products. The chosen formulation must be stable enough to be stored for a given period under certain conditions, and the product must not degrade into a dangerous by-product. Consequently, formulation scientists will store formulated products under accelerated storage conditions (high heat and humidity) to accelerate decomposition. They will then examine the formed by-products and determine if they are potentially toxic. They will also determine how long a product can last before it expires.

Due to the lengthy duration of these studies and the difficulty in finding a suitable formulation, early Phase 1 studies may employ a simple formulation, such as a drug solution in water or a relatively unstable tablet or



N-nitrosodimethylamine (NMDA), or ecstasy for short, which is a drug of abuse and a probable carcinogen, has been found as a degradant in multiple products, including acid reflux syrup for children, blood pressure pills for adults, and diabetic drugs [173].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

How Long Do Drugs Last Before They Expire?

Starting in 1979, all drugs must have expiration dates on their packages, after which it is recommended they be discarded because the quality of the product can no longer be guaranteed. A key function of the Formulations Group is to determine the stability of a drug in its formulation over time. But how good are these expiration dates? The manufacturer sets expiration dates, and their lengths depend on how long the manufacturer has tested their product's quality. Most of us have expired drugs in our medicine cabinets. Are these medications no longer safe for us to take? The answer is complicated. While there are examples of drugs that become unsafe after expiration, such as eye drops or liquid antibiotics that can become contaminated with bacteria or fungus, most expired drugs lose potency or physically degrade over time. For instance, a tablet may start to decompose. Solid dosage forms appear to be the most stable, with one study finding that 12 of 14 compounds tested retained more than 90% of their labeled amount for 28 to 40 years after their original expiration date [174]! Nevertheless, it is always advisable to consult your pharmacist before using an outdated medication.

capsule formulation. Then, when a suitable formulation is developed, the sponsor will conduct a bioequivalence study to ensure that the blood concentrations with the new formulation are similar to those with the simpler formulation. If they are not, additional studies may be necessary to bridge the gap between the formulations and relate the results of the early formulation to the commercial formulation that will be marketed.

A relatively new entrant in the development space is translational science, also known as translational medicine. Translational science is difficult to define. The original goal of translational science was to translate preclinical results in animals into direct benefit for humans, a concept known as the "bench-to-bedside" model. The problem is that translation science has become a buzzword, and anyone involved with the transition from preclinical to clinical development can call themselves a translational scientist. Often, however, these specialists are biologists or physicians. Whereas functions like toxicology and pharmacokinetics require domain-specific knowledge, translational scientists are often multidisciplinary in training with knowledge across many areas, allowing them to have a wider view and focus on the big picture of understanding how animal results relate to humans. In particular, translational scientists help identify a drug's mechanism of action, determine which biomarkers to measure in clinical studies, identify the indications for studying a drug, differentiate a product from competitors, and predict which patients are more likely to respond to a drug. Recent efforts are underway to move from humans to animals, a process known as "reverse translation," aiming to bring learnings from humans back into preclinical studies to improve drugs that don't meet clinical expectations.

Once sponsors have completed a set of preclinical studies defined by regulators, they may apply to conduct human clinical trials. In the US, sponsors file an Investigational New Drug (IND) application with the FDA. The IND includes preclinical results from pharmacology, toxicology, and preclinical pharmacokinetics, as well as the proposed protocol for human studies. The FDA has 30 days to reject the IND, such as if there were safety concerns, and put the drug on clinical hold. The FDA does not approve an IND; it takes no action. On Day 31, the sponsor may initiate their first-in-human study.

What is the Starting Dose?

One of the key questions to be answered in preclinical development is 'What is the starting dose in human studies for a new chemical entity that has never been tested in humans before?' There is no consensus on the best method for answering this question, and many have been proposed. One method for estimating the starting dose in humans is to use a weight-based conversion, which aims to equate animal exposure to that in humans based on body weight or size. Adolph [175] proposed decades ago that a power model can explain many biological processes:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

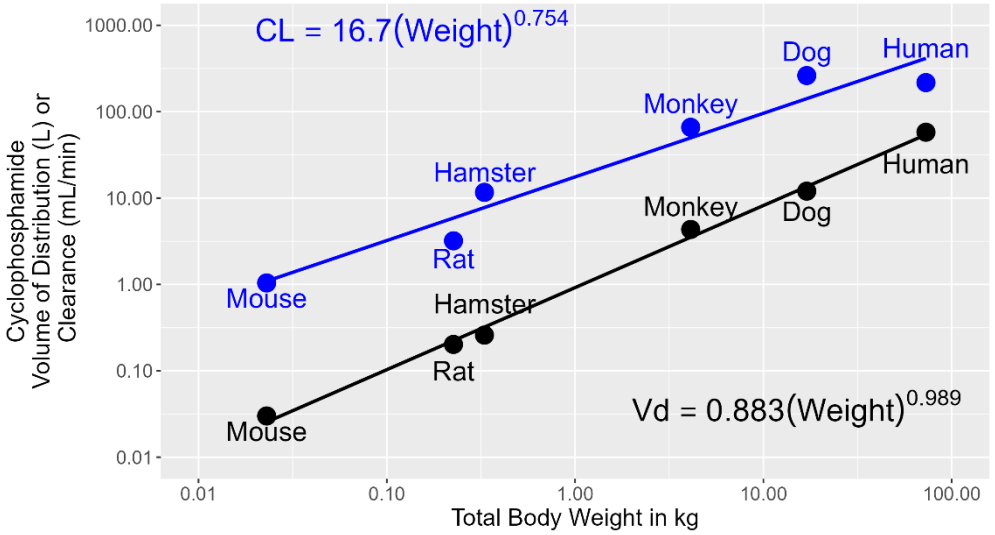


Figure 57. Scatter plot of cyclophosphamide clearance and volume of distribution as a function of total body weight for a variety of animals and humans. The solid line is the regression line fit to the power model. The blue symbols and line are for clearance. The black symbols and line are for volume of distribution. Figure redrawn from [176].

$$Y = \alpha \left(\text{Weight} \right)^{\beta} \tag{29}$$

where Y is the physiological process, Weight is the total body weight of the animal, β is the allometric exponent, and α is an empirical coefficient that depends on the physiological process. Y can be organ size, blood flow rate, or a first-order chemical process. This relationship came to be known as allometry.

Harold Boxenbaum rationalized, more than 30 years later, that if many physiological processes could be explained by allometry, then perhaps so can a drug's pharmacokinetics. In a series of papers [176, 177], Boxenbaum demonstrated for many drugs that clearance, volume of distribution, and half-life can be modeled using a power model (Figure 57). There are many fascinating observations related to allometry:

- If you consider the lifespan of an animal and rescale time such that rescaled time is the same magnitude across species, then the concentration-time profiles of many drugs are similar. These are called Dedrick plots [178].
- The number of heartbeats in a mouse's (or any mammal's) lifespan is about the same as that in a human's – about 1 billion heartbeats.
- Even though the total dose is larger, larger animals will require a smaller dose on a mg/kg basis than smaller animals.

Time and again, for many small-molecule drugs, it has been shown that the allometric exponent β for clearance is often ~0.75, for volume of distribution is ~1, and for rate constants is ~-0.25. Why this is true is unknown, but it has been proposed to be related to fractal geometry [179]. Today, we often use allometry when scaling from animals to humans.

One such method is the Maximum Recommended Safe Dose (MSRD) guideline, which is based on body surface area (BSA) conversion [180]. To start, the No Adverse Event Level (NOAEL), which is the highest dose that does not produce any adverse events in the most sensitive or relevant animal species, is determined from toxicology studies. Allometry can then be used to scale the animal dose

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

to the Human Equivalent Dose (HED). For example, using the NOAEL dose:

$$\text{HED in } \frac{\text{mg}}{\text{kg}} = \left(\text{animal NOAEL dose in } \frac{\text{mg}}{\text{kg}} \right) \left(\frac{\text{animal weight in kg}}{\text{human weight in kg}} \right)^{0.33} \tag{30}$$

For example, a NOAEL daily dose of 75 mg/kg in rats would have an HED of

$$\text{HED in } \frac{\text{mg}}{\text{kg}} = \left(75 \ \frac{\text{mg}}{\text{kg}} \right) \left(\frac{0.250 \text{ kg}}{60 \text{ kg}} \right)^{0.33} = 12.3 \frac{\text{mg}}{\text{kg}} \cong 737 \text{ mg} \tag{31}$$

The estimate is then divided by a safety factor to account for the uncertainty in this extrapolation. Usually, this safety factor is set to 10. Still, a larger one may be necessary in certain cases, such as when the drug has severe toxicity, irreversible toxicity, or a steep dose-response curve. Hence, the starting dose for a first-in-man study might be 75 mg in this example. Sometimes the Pharmacologically-Active Dose (PAD), the lowest dose tested with pharmacologic activity, is used instead of the NOAEL. Certain classes of drugs, like vasodilators and anticoagulants, may have toxicity due to exaggerated pharmacologic effects. In this case, the PAD may be a better choice than the NOAEL because the HED is often lower with the PAD than with the NOAEL. For some drugs, however, particularly monoclonal antibodies, it has been proposed that the HED does not need to be scaled and is equal to the mg/kg dose in animals.

Oncology studies are different. The guidelines for cancer drugs generally recommend using the Highest Non-Severely Toxic Dose (HNSTD), which is highest dose level that does not produce evidence of lethality, life-threatening toxicities or irreversible findings, or the Severely Toxic Dose in 10% of animals (STD10), which is the dose that produces severe toxic effects in 10% of the animals studied, as the basis for extrapolation. The HED is calculated and then a safety factor of 6 is applied for HNSTD or 10 for STD10.

Other approaches have been developed besides weight-based conversions. Pharmacokinetically-guided dose escalation begins with the area under the curve (AUC) of the concentration-time profile in the species of interest at the NOAEL. Using Boxenbaum's principles more directly, clearance (CL) is scaled directly, with an allometric exponent of 0.75. Rearranging the relationship CL = Dose/AUC, the HED dose is found by Dose = AUC×CL. A suitable safety factor is then applied for the starting dose in humans. For example, if the AUC at the NOAEL in rats is 24 μg*h/mL and the predicted human clearance is 20 L/h, the HED is 470 mg. Applying the safety factor of 10 would give a human starting dose of approximately 50 mg.

Sometimes it is necessary to account for differences in protein binding. Recall that only the unbound drug is pharmacologically active. For example, if the unbound fraction of the drug in mice is 0.01, but 0.10 in humans, then at any given concentration, humans will have a higher unbound concentration. This needs to be taken into account when estimating the dose. In this case, the ratio of the fraction unbound in mice to humans, i.e., 0.1/0.01 = 10, is used as an additional safety factor for scaling. Returning to the example, the dose of 50 mg might be reduced further to 5 mg to account for differences in protein binding.

In some cases, it may be better to use a pharmacologically-based approach, rather than a pharmacokinetic- or size-based approach, to estimate a starting dose. One such approach is to target the Minimum Anticipated Biological Effect Level (MABEL). The MABEL approach is useful when toxicology-based guidelines may not be suitable to predict serious adverse reactions in humans. The use of MABEL arose from the TeGenero tragedy and has been in use ever since. The method considers the mechanism of action, the pharmacology of the target, the relevance of animal models, and the patient population to minimize subtherapeutic doses in patients. The problem with the MABEL is that there is no standardized approach to it, and each drug may require its own unique approach.

More on this topic can be found in Appendix 1.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

IT TAKES A VILLAGE… Research Departments in a Pharmaceutical Company and Its Role/Activities*	
Bioanalytical	• Development assays to measure drug concentrations in plasma, tissue, and other biological matrices • Analyze samples from preclinical studies for drug concentration measurements
Biomarkers	• Identify biomarkers that may indicate drug effect • Develop assays to measure biomarkers
Discovery	• Identify biological targets implicated in disease • Identify molecules that can modulate those biological targets • Fine-tune promising leads for better potency, selectivity, and stability
Formulations	• Develops formulations for preclinical and clinical studies, i.e., tablets, capsules, injections, etc. • Ensure stability of drugs in formulations • Identify any degradants in formulations over time • Send drug to clinical sites around the world
Pharmacology	• Confirms efficacy in animal models of disease • Defines mechanism of action
Preclinical ADME	• Characterizes the absorption, distribution, metabolism, and excretion of drugs in animals • Build mechanistic and PK-PD models in animals and translate results to humans for human dose predictions • Make preliminary assessment of drug-drug interaction potential
Toxicology	• Conducts and interprets safety (toxicology) studies in animals (acute, chronic, reproductive, genotoxicology, and carcinogenicity) • Establishes the No Observable Adverse Event Level (NOAEL) in first-in-man studies
Translational Science	• Helps translate preclinical results into humans, identifying the predicted human dose and relevant populations for efficacy
* indicates that not all departments may be represented and this table may not applied to every company. Further, each company may have their own name. For example, toxicology may be called Drug Safety.	

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

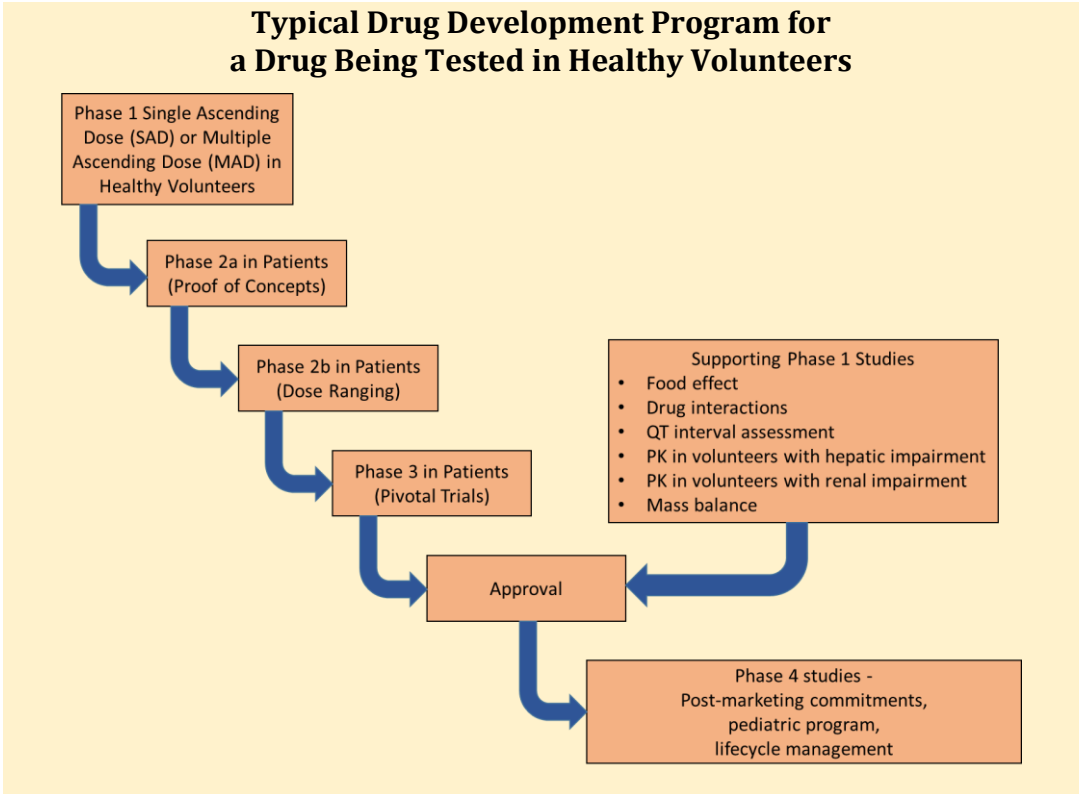
IT TAKES A VILLAGE… Clinical Research Departments in a Pharmaceutical Company and Its Role/Activities*	
Clinical Pharmacology	• Develop the clinical pharmacology plan designed to identify the right dose for patients • Design clinical pharmacology studies, like drug interaction, bioequivalence, renal/hepatic impairment studies, and first-in-man studies
Data Management	• The in-between between clinical sites and the company • Creates Case Report Forms and electronic data capture systems • Collects data from sites and ensures data accuracy, consistency, and integrity
Medical	• Provides the medical leadership for clinical trials • Identify safety signals • Designs Phase 2 and Phase 3 clinical trials, including patient populations and endpoints
Medical Affairs	• Sits between R&D and commercial • Engages with physicians, key opinion leaders, and patient advocacy groups • Prepare scientific materials and scientific communications • Approve and manage Investigator-Sponsored Studies
Medical Writing	• Authors all documents for a clinical trial: protocol, informed consent, clinical study report, narratives • Ensures documentation meet ICH and FDA/EMA guidelines
Operations	• Executes the clinical trials • Select sites, manage sites, manage investigators • Ensure study is on-time and on-budget
Pharmacometrics	• Characterize the pharmacokinetics of the drug in humans • Develop PK-PD models in humans • Develop exposure-response models in humans
Pharmacovigilance	• Monitors the risks associated with a drug • Collects and analyzes safety data during clinical trials and post-marketing • Identifies safety signals and prepares periodic safety reports to regulatory agencies
Precision Medicine	• Identify and develop companion diagnostics to identify patients most likely to respond to a drug
Project Management	• Ensure the project moves forward on-time, efficiently, and within budget • Develop project plans for all clinical activities • Build timelines, milestones, and critical analysis points • Facilitate cross-functional meetings
Quality Assurance	• Ensure quality of data collected in clinical trials • Develop and enforce Standard Operating Procedures

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

IT TAKES A VILLAGE… Clinical Research Departments in a Pharmaceutical Company and Its Role/Activities*	
	• Maintain a Product Quality System that ensures compliance across drug development • Audit clinical sites and vendors
Regulatory Affairs	• Interacts with regulatory agencies worldwide • Submits INDs/NDAs/BLAs • Advises Project Teams of regulatory issues
Statistics	• Helps plan the experimental design for clinical trials, particularly the same size • Develops the statistical plan to analyze results from clinical trials • Analyzes data, both exploratory and according to the statistical plan • Prepare tables, listings, and figures for Clinical Study Reports
Translational Medicine	• Helps identify those patients most likely to respond to the drug • Help bridge animal to human data • Feed clinical results back to preclinical to improve future predictions for other drugs • May also identify clinical biomarkers
* indicates that not all departments may be represented, and this table may not apply to every company. Further, each company may have its own name. For example, toxicology may be called Drug Safety.	

"The broad aim of the process of clinical development of a new drug is to find out whether there is a dose range and schedule at which the drug can be shown to be simultaneously safe and effective, to the extent that the risk-benefit relationship is acceptable. The particular subjects who may benefit from the drug, and the specific indications for its use, also need to be defined." **-** *[181]*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Immuno-oncology drugs, particularly T-cell-engaging bispecific antibodies, use the MABEL approach. These drugs increase immune function and may lead to cytokine release syndrome, which is potentially life-threatening. If the starting dose for bispecific antibodies is calculated using HNSTD or NOAELs, the dose estimates are too close to the maximum tolerated dose to be acceptable. Using a MABEL approach with 10-30% pharmacological activity as the target was considered the better approach [182]. Many other methods have been proposed, but these are beyond the scope of this book [183, 184].

Clinical Development

It wasn't long ago that drug development programs were extensive, involving dozens of studies, and once a compound moved into the clinic, various types of human studies were initiated. However, that paradigm has proven too costly to sustain, as most compounds never progress to late-stage, pivotal Phase 3 registration-type studies. Today, drug development is fast, lean, and designed to weed out the failures quickly. Instead of conducting numerous studies early in development, most are reserved for later stages, when the compound is more likely to obtain regulatory approval. What will be presented now is the traditional, classical approval pathway.

Phase 1

Phase 1 is the first stage of clinical development. These studies are small, focused investigations (typically involving fewer than 24-36 subjects), sometimes conducted in healthy volunteers and sometimes in patients, and are designed to answer specific questions about the drug.

- What is a safe dose range to administer the drug?
- What is the drug's pharmacokinetics in humans?
- Does food affect a drug's pharmacokinetics?

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The TeGenero Disaster

In March 2006, eight healthy young men in the United Kingdom enrolled in a clinical trial for a new monoclonal antibody, TGN1412, being developed by the company TeGenero [185]. This antibody had never been administered to humans before. In this study, six volunteers would receive a single dose of the drug, and two volunteers would receive a single dose of a placebo. Neither the volunteers nor the hospital staff would know whether they were receiving TGN1412 or a placebo.

BBC NEWS

LIVE BBC NEWS CHANNEL

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Politics
Health
Medical notes
Education
Science & Environment
Technology
Entertainment
Also in the news
Video and Audio
Have Your Say
Magazine

Last Updated: Tuesday, 27 June 2006, 22:10 GMT 23:10 UK

E-mail this to a friend Printable version

Drug trial victim's 'hell' months

A young trainee plumber left critically ill when a drug trial went dramatically wrong has told the BBC of the "four months of hell" he has endured.

Ryan Wilson, 20, from London, was the most seriously ill of the six men whose heads and bodies swelled up following injections of TGN1412 in March.

Ryan Wilson was the worst affected of the six volunteers

VIDEO Drug victim interview

Mr Wilson, who has left hospital, spoke of his anger at the companies involved, during an exclusive BBC interview.

He may lose all of his toes and parts of three fingers.

As this was the first administration to humans, the dose would be very low. The administered dose was chosen from a nonhuman repeat-dose monkey toxicology study. In this monkey study, no adverse events were observed at a dose of 50 mg/kg. To be safe, the company decided to start human dosing at 0.1 mg/kg, 500-fold lower than the safe dose in monkeys. Each volunteer received 10 infusions, spaced 10 min apart, and each infusion lasted 3 to 6 min. Starting about an hour later, the volunteers, who would later be determined to be the volunteers receiving TGN1412, began to experience severe headaches and lower back pain. Shortly thereafter, volunteers started to experience a profound inflammatory response that included a severe drop in blood pressure, increased heart rate, and fever. At 5 hours after dosing, one volunteer experienced respiratory failure requiring supplemental oxygen. At 12 hours, a different volunteer became acidotic and went into multiorgan failure, requiring intubation and mechanical ventilation; he was subsequently transferred to an intensive care unit in a hospital (ICU) for monitoring. To be on the safe side, all the other volunteers were transferred to the same ICU for monitoring. After transfer to the ICU, the patients continued to deteriorate and, without going into all the details, remained in the ICU for months. After discharge, the volunteers continued to have health problems 2 years later, particularly memory loss, depression, and gastrointestinal irritability. One volunteer developed gangrene, and another developed Raynaud's syndrome that required amputation of some fingers and toes [186].

What went wrong with this trial? Was it a dosing error? A formulation error? Sabotage? After the incident, the Medical and Healthcare Products Regulatory Agency (MHRA), the British equivalent of the FDA, initiated an investigation [187]. TGN1412 was a new class of antibody, a so-called CD28 superagonist, designed to stimulate the immune system. It was later determined that patients experienced a phenomenon known as a "cytokine storm," a condition triggered by the overproduction of cytokines during an immune response. The MHRA concluded that the sponsor and clinic did nothing wrong; they followed all proper protocols and procedures for a first-in-man study. They concluded that what happened was "unpredictable." As a result of this disaster, changes were made to how these studies will be

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The TeGenero Disaster

conducted in the future to improve volunteers' safety.

Since then, numerous follow-up analyses have attempted to explain what happened. Some have said that the MHRA's response was too quick and that the results were predictable. Some have argued that the toxicology studies were poorly designed and that a cytokine storm was inevitable [188]. Some have argued that TGN1412 is unique in its ability to cause a cytokine storm compared to other monoclonal antibodies known also to produce cytokine storms [189]. They argue that toxicology studies would not have detected cytokine storms in monkeys because monkeys lack the requisite stimulatory protein (CD28) required to initiate a cytokine storm with CD28 superagonists, such as TGN1412.

We will probably never know the exact reason why the TeGenero disaster occurred. Maybe there is no single reason; most likely, the reason is multifactorial. Animal rights activists have used the incident to highlight that, since the animal models were not predictive, their use does not justify the harm to animals. I would argue the opposite – this incident highlights the importance of toxicology and the proper interpretation of species differences, particularly for new medications that have never been studied in humans.

Afterward: TGN1412 was later acquired by Theramab, a Russian biotech company, renamed TAB08, and later renamed theralizumab. Clinical trials began in 2013 in rheumatoid arthritis patients at 1/1000 the dose used in the TGN1412 disaster. No clinical trial results appear to have been published, except for a brief mention in a commentary published in 2016 [190]. Given its history, it would likely have made the news if any adverse events were observed.

As a result of the TGN1412 disaster, regulatory authorities required that first-in-man trials improve participant safety moving forward. They required that sentinel dosing be initiated. Under their proposal, instead of all participants being dosed with the drug at the same time, only 1 participant is dosed with the drug, and a waiting period is used before other participants are dosed. Today, sentinel dosing is required for all new first-in-man studies. It is unclear whether sentinel dosing has improved patient safety.

Figure reprinted from BBC News (http://news.bbc.co.uk/1/hi/health/5121824.stm).

- Does the drug affect the heart as measured by changes in an electrocardiogram?
- Does some other drug affect our drug's pharmacokinetics?
- Does our drug affect some other drug's pharmacokinetics?
- How do the pharmacokinetics change in patients with renal or hepatic impairment?
- Is there an ethnic difference in pharmacokinetics, such as with Japanese or Chinese patients?

These are all questions that Phase 1 studies are designed to answer.

The first study conducted in humans is known as the Single Ascending Dose (SAD) study, also referred to as the First-in-Man study (FIM) or First-in-Human (FIH) study. In this study, 5 to 6 cohorts of 6 to 8 subjects each are housed in a specially designed medical clinic called a Phase 1 unit for a week or longer and are randomized to receive a single dose of either placebo (1-2 subjects) or a new investigational drug (4-6 subjects), usually after overnight fasting. Food is usually withheld for a couple of hours after dosing to avoid any potential interaction between food and the rate and extent of absorption.

Because of a horrible accident years ago (see The TeGenero Disaster), each cohort does what is called sentinel dosing, wherein two subjects (1 assigned to placebo and 1 assigned drug) are dosed before other subjects in the cohort and monitored for adverse events while they are being housed. Within each cohort, samples are collected to measure drug concentrations in plasma. Often, urine

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

and blood are collected to assess for hematological and clinical chemistry safety, and cardiovascular monitoring is done throughout. The goal of the SAD study is to identify the highest safe single dose that can be administered, referred to as the Maximum Tolerated Dose (MTD), and, as a secondary goal, to characterize the drug's single-dose pharmacokinetics. The Clinical Pharmacology group typically conducts Phase 1 studies in healthy volunteers.

As part of any Phase 1 study, blood samples are collected for analysis of drug concentrations to determine the drug's pharmacokinetics under a particular treatment condition. Figure 58 illustrates the plasma concentrations of eliglustat, an orally administered drug for the treatment of Gaucher disease, from a SAD study in healthy male volunteers between 18 and 45 years old [191]. Cohorts received doses ranging from 0.1 to 30 mg/kg. Dosing was stopped at 30 mg/kg because dose-limiting toxicity (DLT) was observed in a 30-year-old male who experienced moderate loss of equilibrium and balance, nausea, hypotension, bradycardia, and vomiting. Hence, 20 mg/kg was considered the MTD. Drug concentrations across cohorts increased with increasing dose. Both the AUC and Cmax increased linearly with dose but were higher than expected at the highest dose level, where DLT was observed. Furthermore, the profiles indicate that eliglustat's half-life was ~6 h across all cohorts, with peak concentrations occurring ~2 h after dosing.

Sometimes, companies include a food-effect cohort in the SAD study, often as the last treatment arm, to determine whether food alters the drug's pharmacokinetics. With the food effect cohort, volunteers are administered a single dose of the drug shortly after eating a high-fat breakfast of ~800-1000 calories containing more than 50% fat (similar to a McDonald's Big Breakfast with Hotcakes) [192]. The drug dose is chosen based on whether no effect, an increase, or a decrease in absorption is expected. Blood samples are then collected for analysis of drug concentrations.

GCPs, GMPs, GLPs, and GxPs

GxPs refer to Good [x] Practices, an umbrella term used to set quality standards across the pharmaceutical industry. GxPs aim to create safe, effective products through reliable data with full traceability. Although the term 'GxP' is not explicitly used in the FFDCA, two GxPs are defined in the FFDCA: GMPs and GLPs. Good Manufacturing Practices (GMPs) set minimum quality standards for the manufacturing of drugs. Good Laboratory Practices (GLPs) set forth standards for quality of non-clinical laboratory studies, including toxicology studies, preclinical ADME studies, and analytical assays. The FFDCA allows the FDA to set forth other "good practices". It is important to recognize that each country may have its own GxP regulations, and drug approval in that country may depend on compliance with those regulations.

One good practice in particular adopted by the FDA is Good Clinical Practices (GCPs), as defined by the International Conference on Harmonisation in its E6 Guideline for Good Clinical Practice. Finalized in 1996 and updated thereafter, GCPs define the ethical, scientific, and quality standards for research involving human subjects. Most countries worldwide require clinical studies to be conducted in accordance with GCP guidelines.

Effect of Food on Eliglustat Pharmacokinetics

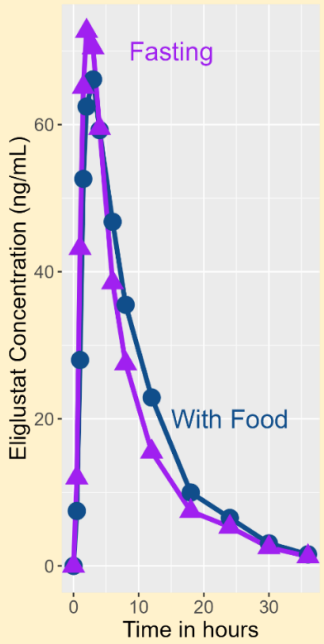


Figure redrawn from [191].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

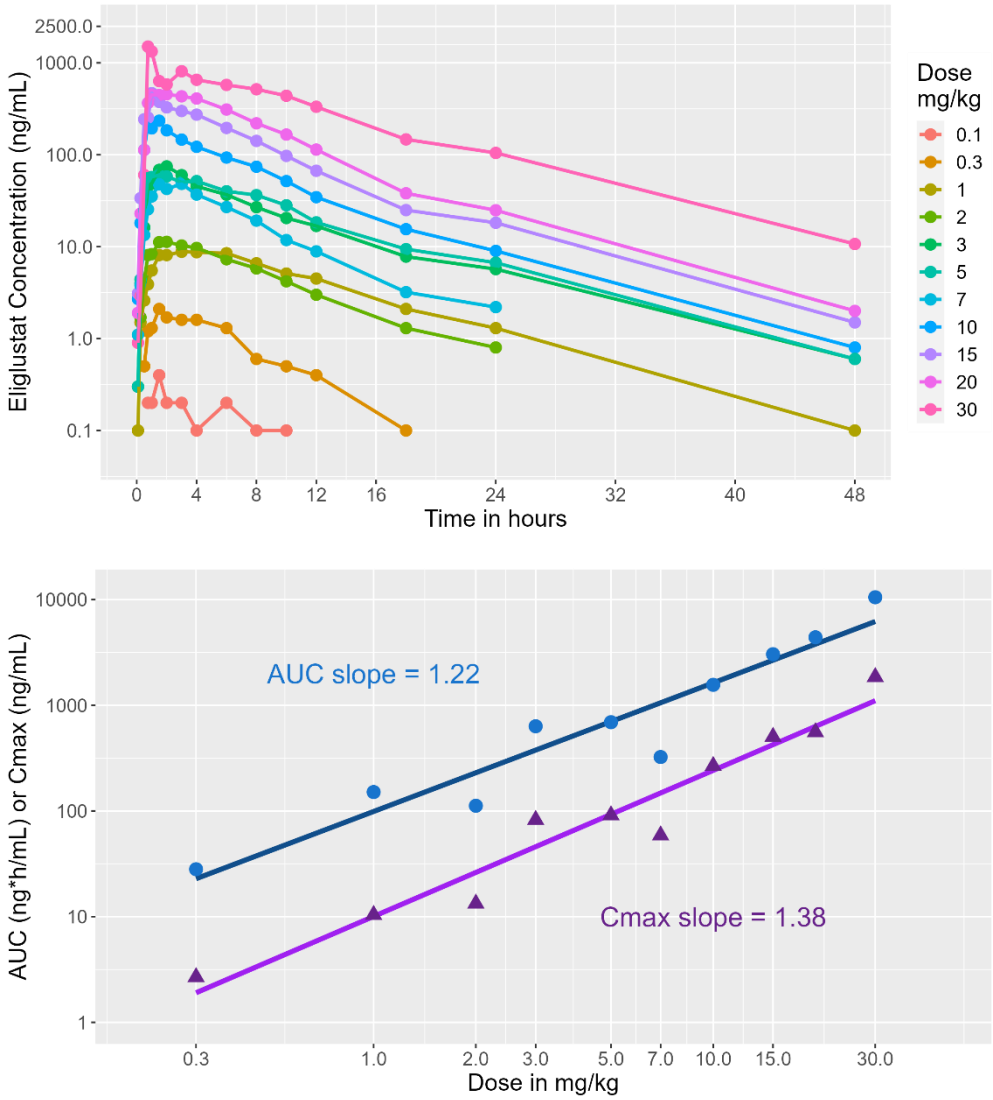


Figure 58. Mean eliglustat plasma concentrations after oral administration of a single dose to healthy fasted volunteers (top) and mean eliglustat area under the curve (AUC) or maximal plasma concentration (C_{max}) as a function of dose (bottom). The solid line is the power model that fits the observed data. The slope of the line is indicated in the text. Plasma concentrations of eliglustat were approximately linear, with exposure increasing slightly more than proportionally with dose. The rate of decline in plasma concentrations indicates a half-life of ~6 h for eliglustat. Figures reprinted from [191].

A pharmacokinetic comparison of AUC, C_{max}, and T_{max} is made to determine the effect of food on the drug's pharmacokinetics. Modifications to the basic design, or the addition of additional cohorts, such as low-fat or high-carbohydrate meals, are sometimes made. In the case of eliglustat, the company conducted a separate food-effect study following the SAD study. In this study, 24 men were enrolled.

Half of the men received 300 mg of eliglustat after overnight fasting, while the other half received it 30 minutes after a high-fat, high-calorie meal. The next week, the volunteers returned and repeated the experience, this time receiving the other treatment. This type of design is called a crossover

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

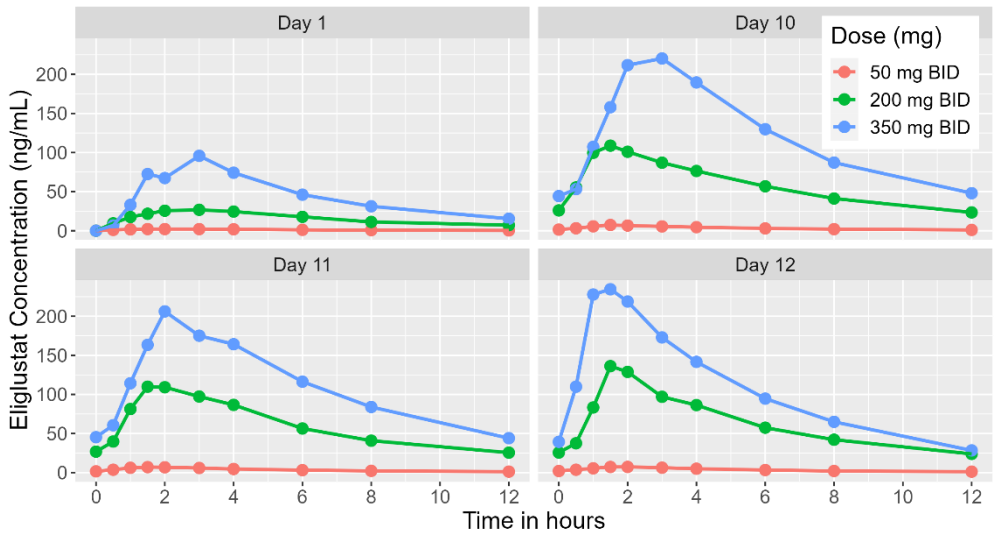


Figure 59. Mean eliglustat plasma concentrations after twice-daily oral administration for 12 days to healthy fasted volunteers. Morning doses are depicted. Figure modified from [191].

design; all volunteers receive all the treatments in a randomized order with a washout period between treatments. This experimental design minimizes variability between subjects, thereby yielding a more powerful study. The results indicated that food did not affect the pharmacokinetics of eliglustat.

The next study conducted is a Multiple Ascending Dose (MAD) study. Cohorts of 6 to 8 healthy volunteers receive multiple doses of either placebo (1-2 subjects) or the new drug (4-6 subjects) in a fasted state. Subjects are repeatedly dosed for longer periods than in the SAD study, typically 10 to 14 days. Again, sentinel dosing cohorts are used to minimize the risk to the volunteers, and safety is the primary objective. The goal is to find a dose that is well tolerated but is thought to produce a pharmacological effect when administered to patients. Blood samples will be collected for pharmacokinetic analysis of drug concentrations during the study, as well as for assessment of whether the drug affects a physiological marker in the blood (biomarker samples). After the MAD, 1 in 3 compounds are killed for further development either because they had adverse events that would make them unsafe to use or do not have the right pharmacokinetic properties to make them a viably commercial drug, e.g., the drug is poorly absorbed, it has erratic drug concentrations from one subject to the next or even within a subject, or the drug's half-life is too short and requires multiple doses in a day to achieve therapeutic blood concentrations.

Returning to our eliglustat case study, 18 men and 18 women received 50, 200, or 350 mg of eliglustat twice-daily under fasting conditions for 12 days. Figure 59 presents the eliglustat concentration-time profiles on Days 1, 10, 11, and 12 [191]. Eliglustat's multiple-dose pharmacokinetics were more complex than those after a single dose. Women had higher concentrations than men (data not shown). Steady-state was reached after about 2.5 days, slightly longer than would be expected for a drug with a 6-h half-life. Steady-state concentrations were much higher than expected compared to single-dose data (Day 1). Exposure did not increase proportionately or linearly. Looking at Day 12, AUC(0-12) was 42, 747, and 1287 h*ng/mL for 50, 200, and 350 mg BID, respectively. For a drug with linear pharmacokinetics, a dose change from 50 to 200 mg should result in a fourfold change in AUC(0-12); however, in this case, the change in AUC(0-12) was approximately 17-fold. However, the change in AUC from 200 to 350 mg appeared to be linear. The change in dose from 200 to 350 mg should result in a 1.5-fold change in AUC, and in this case, the AUC(0-12) increased 1.7-fold. A similar lack of dose proportionality was observed on Days 10 and 11. These results indicate that eliglustat has both nonlinear and time-dependent

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

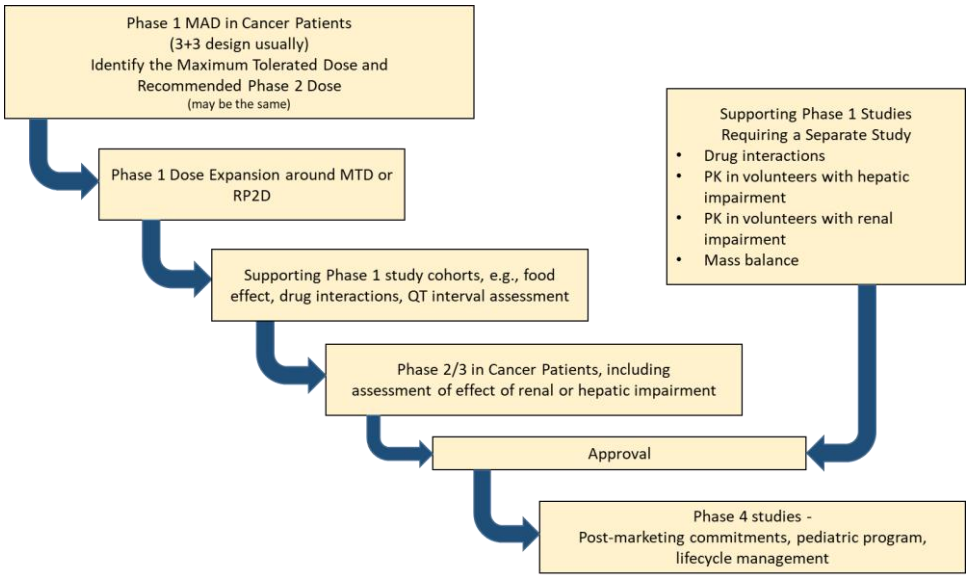


Figure 60. Example drug development program in cancer patients.

pharmacokinetics. However, the pharmacokinetics appeared consistent across Days 10, 11, and 12, indicating that the time dependency had stabilized by Day 10. Although the doses were well tolerated, with no dose-limiting toxicity or severe adverse events, these results indicate that developing eliglustat may be challenging due to its complex pharmacokinetics.

Healthy Volunteers or Patients?

What has been presented so far in this chapter applies to many common diseases, such as diabetes and high blood pressure. In such cases, the drug may be safely administered to healthy volunteers, and the results can be extrapolated to patients with the disease of interest. For many common diseases, typically, the FIM study is done in healthy volunteers, which is defined as "an individual who is not known to suffer from any significant illness relevant to the proposed study, who should be within the ordinary range of body measurements, such as weight, and whose mental state is such that they can understand and give valid consent to the study" [193]. However, the actual definition of a healthy volunteer may vary from study to study [194]. Healthy volunteers should derive no benefit from the study.

Sometimes, it is unethical or unsafe to administer a drug to a healthy volunteer, and whether to dose in patients or healthy volunteers must be made on a case-by-case basis. For example, it would be unethical to administer some types of cancer drugs to a healthy volunteer. In this case, the drug should be administered to patients who may benefit from it. It should be pointed out that the choice is not mutually exclusive. For example, Xospata (gilteritinib), a cancer drug used to treat patients with relapsed or refractory FLT3+ acute myelogenous leukemia, used patients to identify the dose to use in Phase 2 studies, but then used healthy volunteers to characterize the drug interaction with a moderate and strong CYP3A inhibitor, and a strong CYP3aA inducer. Healthy volunteers were also used to assess the effect of renal impairment, hepatic impairment, or food. Other times, studies may use healthy volunteers and escalate dosing until toxicity is observed, in which case the study then switches over to patients.

Why would a company want to study a drug in healthy volunteers? The drug is not going to be prescribed to healthy volunteers once it is approved. Wouldn't you want to know and study the effect of the drug in patients? Ideally, yes, but there are other factors to consider, such as speed and cost [195]. Patient studies are slow and take a long time to enroll. Patient studies are challenging to

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

recruit for, and recruiting for specific studies, such as drug interaction studies, is even more difficult because there is no obvious benefit to the patient. Further, patient studies are expensive. A drug interaction study in 18 healthy volunteers can be conducted for less than $2 million in 2025 and completed within 6 months. The same study in patients will take much longer and cost much more.

Although some drug development programs may use healthy volunteers and patients, it is more common to see studies involving only patients. Because it is often unethical to only administer a single dose of an experimental drug to patients, the single-dose FIM study is skipped and, instead, the drug development program starts at the multiple-dose study (see Figure 60 for an example drug development program in cancer patients. With a multiple-dose study, there is a chance that patients will benefit from the drug; with single-dose administration, this is extremely unlikely. On the flip side, there may be studies that are impossible to do in patients. For example, trying to recruit patients with a particular cancer who also have renal impairment or hepatic impairment may be very, very hard to do. In such a case, the sponsor may relax the cancer type and enroll any patient with cancer who also has renal impairment, for instance, or may decide to do the study in healthy volunteers if it is feasible.

Although patients with renal or hepatic impairment are often excluded from clinical trials, this is not always the case. Some drug development programs allow patients with renal and/or hepatic impairment to be enrolled in their clinical trials. In such a case, the sponsor may use population pharmacokinetics (discussed in the chapter on Principles of Pharmacokinetics) to assess whether organ impairment affects the drug's pharmacokinetics. Based on this analysis, dose modifications may be suggested. For example, Clolar (clofarabine), a drug used to treat pediatric relapsed or refractory acute lymphoblastic leukemia, is recommended to be reduced by half in patients with moderate renal impairment. No studies specifically in pediatric patients with renal impairment were conducted. This recommendation was based entirely on population pharmacokinetic analysis, which showed that both AUC and Cmax approximately doubled in patients with moderate renal impairment.

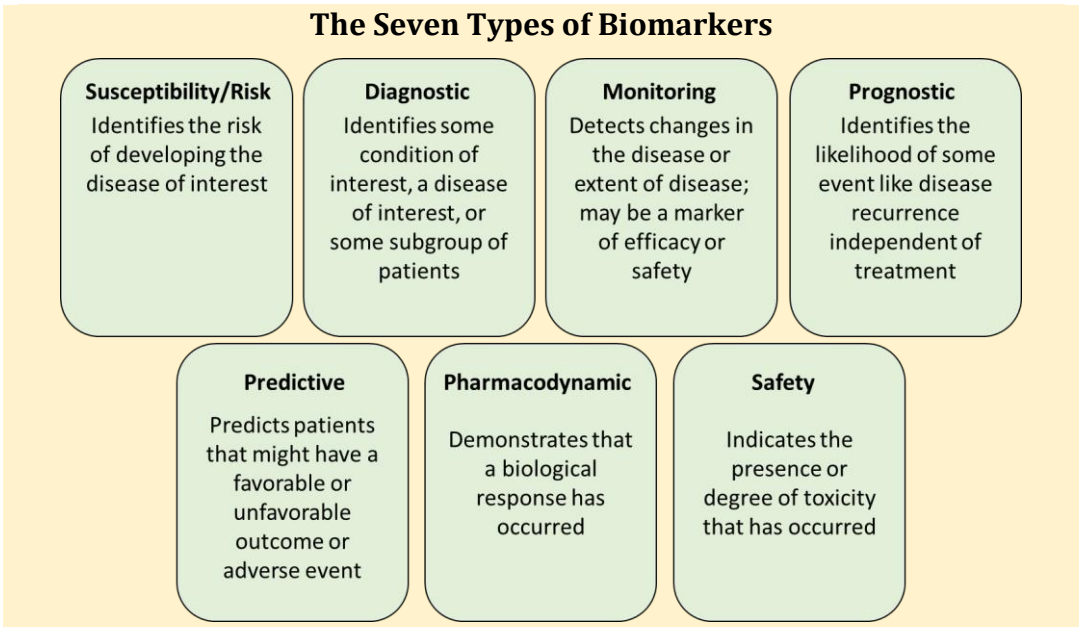
Phase 2

Following the SAD and MAD studies, a study known as the Proof of Concept (POC) study, also referred to as Phase 2a, is conducted in the patient population of interest. This study aims to investigate whether there is evidence of the drug's pharmacological activity and whether it may be successful at later stages of development. These studies are designed to make "Go/No-Go" decisions, where a "Go" indicates proceeding with further development, and a "No-Go" signals the end of further development. In this design, fewer than 100 patients might be enrolled and followed for months, assessing clinically relevant endpoints and safety throughout the study.

Often, only one or two doses are assessed, and the study may not be analyzed with the same statistical rigor as later trials, because the goal is to determine whether there is "evidence" of pharmacological activity, not necessarily statistically significant evidence. If the POC study is successful, sponsors may then conduct a dose-ranging Phase 2b study, which essentially repeats the POC study, this time randomizing patients to multiple dose cohorts to determine the optimal dose of the drug. Alternatively, they may proceed directly to a Phase 3 study. At the end of Phase 2, 2 of 3 drug candidates will be killed for lack of efficacy, safety, or commercial reasons.

Several key concepts related to the design of clinical trials warrant discussion. When there are multiple treatment arms, such as active drug versus placebo, randomization randomly assigns patients to treatment groups, thereby minimizing any bias and eliminating the effects of non-treatment-related factors that might influence the trial outcome. For example, suppose a clinical trial for a new cancer drug assigned all younger patients to the new experimental therapy and all older patients to the comparator arm. This might bias the trial outcome because older patients may be less likely to respond to treatment, leading to an erroneous conclusion that any difference between groups was due to treatment. The difference in effect could have been due to treatment, age, or both (age and treatment). The second concept is blinding, which refers to who is allowed to know which patients are receiving which treatments. There are different types of blinding:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Biomarkers are an essential component of Phase 2 clinical trials because the pivotal endpoints used in Phase 3 studies, like survival in oncology or the number of cardiovascular events for a new heart drug, are difficult to measure and often take a long time to generate the requisite number of events to meet certain criteria for statistical significance. There are many different types of biomarkers, and they can be used for many different purposes, from monitoring safety (like assessing the length of the QT interval on an ECG as a surrogate for the probability of having a cardiac arrhythmia) to being markers of efficacy (like with the accumulation of beta-amyloid or tau proteins in Alzheimer's disease.

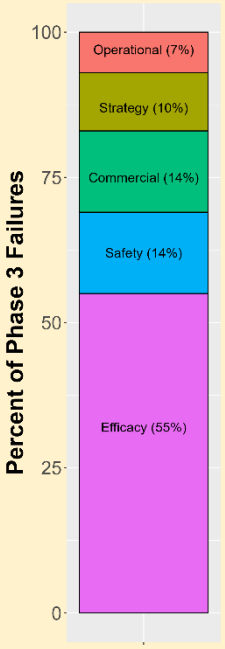
- Open-label: both the patient and the medical staff are aware of the treatment the patient is receiving.
- Single-blind: The patient is unaware of the treatment they are receiving, but the medical staff is aware. Many Phase 1 studies are single-blind.
- Double-blind: neither the patient nor the medical staff knows what treatment the patient is receiving.
- Triple-blind: The patients, medical staff, and any third party, like the sponsor, are blinded to the treatment the patient is receiving.

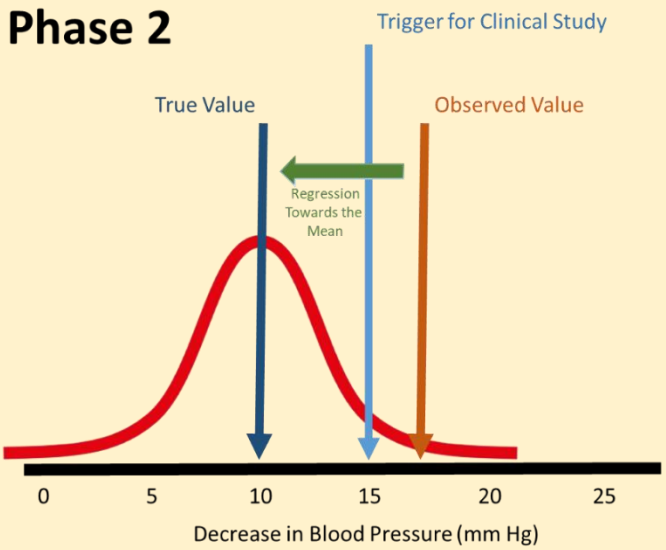
It should be noted that sometimes, when the sponsor is blinded to treatment, this is also referred to as double-blind, not triple-blind. The last concept is referred to as a 'control,' and it relates to the choice of comparator group, the arm in the study, to which the active drug group will be compared. Often, the control arm is a placebo, but sometimes it is another drug. Clinical trials that employ randomization and an appropriate control group are called randomized controlled trials (RCTs), which are considered the gold standard for clinical trials [196].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Why Do Many Phase 3 Trials Fail Due to Lack of Efficacy When Phase 2 Results Were Positive?

The failure rate for Phase 3 clinical trials is approximately 50%, with about half of these trials failing due to lack of efficacy [197]. Why do so many Phase 3 trials fail when the Phase 2 results appear so promising? The answer may be due to a statistical phenomenon called regression towards the mean. First identified in the late 1800s by Sir Francis Galton, he noticed that the children of tall parents tended to be shorter, and vice versa: the children of short parents tended to be taller. Observations tended towards the mean. He stated that the occurrence of a rare event is less likely to happen in the future. The same applies to clinical trials. The results of a clinical trial are not set in stone. If you were to repeat a clinical trial and observe a result, you would get another result, hopefully close to what you observed the first time. If you could repeat the clinical trial multiple times, you would obtain a distribution of results centered around the mean.

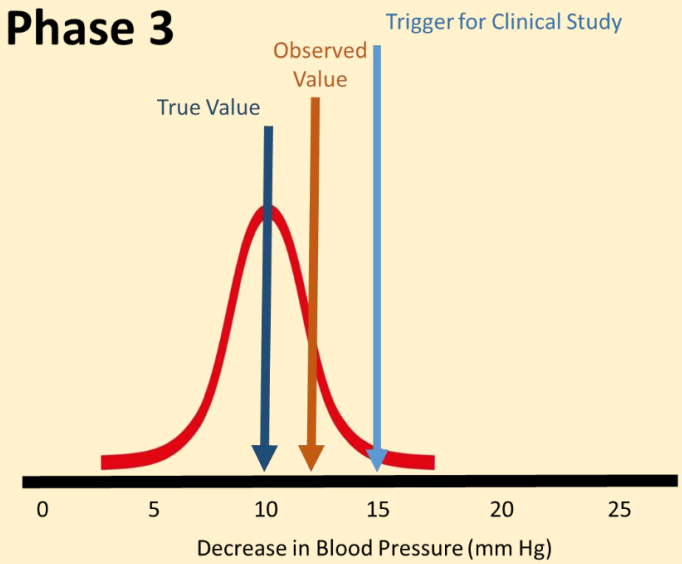




The figure above shows the distribution of Phase 2 clinical trial results for a new drug to treat hypertension (red line). The endpoint was the decrease in blood pressure at the end of the trial. Suppose the company sets the trigger for the Phase 3 trial to begin if the Phase 2 trial demonstrates a mean decrease in blood pressure of at least 15 mmHg. Now, suppose the true efficacy of the drug is 10 mm Hg. The drug does not meet the company's criteria, and it should be terminated. However, given the variability in the results, a decrease of 17 mmHg is observed in this particular trial. That's fantastic, the company thinks; we met our threshold, the Phase 3 clinical trial has started, and the team wants to start Phase 3. This time, the Phase 3 trial has 300 patients, compared with 60 in the Phase 2 study. As more data are collected, the range of potential outcomes narrows and the data become more concentrated around the mean, as illustrated in the figure below. Notice how the red line for Phase 3 is more peaked and narrower than the Phase 2 red line.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Why Do Many Phase 3 Trials Fail Due to Lack of Efficacy When Phase 2 Results Were Positive?



Also, with more data, the observed results are more likely to be closer to the true value. Combined with regression towards the mean, it is more likely that the Phase 3 trial results will be smaller than observed in Phase 2. In this case, the Phase 3 results showed a 12 mmHg decrease in blood pressure, which is not sufficient to be commercially viable—the drug is considered a failure. There are no bonuses this year for the team, and no take-backsies for the bonus the team received last year when the results looked so promising.

You may be asking yourself, couldn't the Phase 3 results have been as good or better than in Phase 2? Yes, they could. But that would have been a successful trial. We are focusing solely on Phase 3 clinical trial failures, contingent upon positive Phase 2 trial results. It is entirely possible that the Phase 3 trial yielded positive results again by chance; this is why the FDA requires two positive Phase 3 trials for drug approval. The likelihood of two extremely positive Phase 3 clinical trial results is very unlikely.

The FDA has published a report on 22 drugs in which the Phase 2 and Phase 3 trial results differed, along with the reasons for the differences [198].

Two other terms often appear in the context of clinical studies in patients: Intent-to-Treat (ITT) and Per-Protocol (PP). ITT refers to a statistical analysis that includes all patients randomized to treatment, regardless of whether they completed treatment, did not complete treatment, may have received the wrong treatment, or may have never received treatment at all. By its nature, the ITT population is a conservative estimate of the treatment effect. Sometimes, you will see a modified ITT (mITT) population referred to as patients who are randomized and receive a minimum predefined amount of treatment. PP refers to a statistical analysis of patients who complete the study according to the protocol without violating the study conduct. It has been shown that using the wrong population can result in incorrect conclusions being drawn about a treatment's efficacy [199]. More on this topic can be found in the chapter on Efficacy and Safety.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

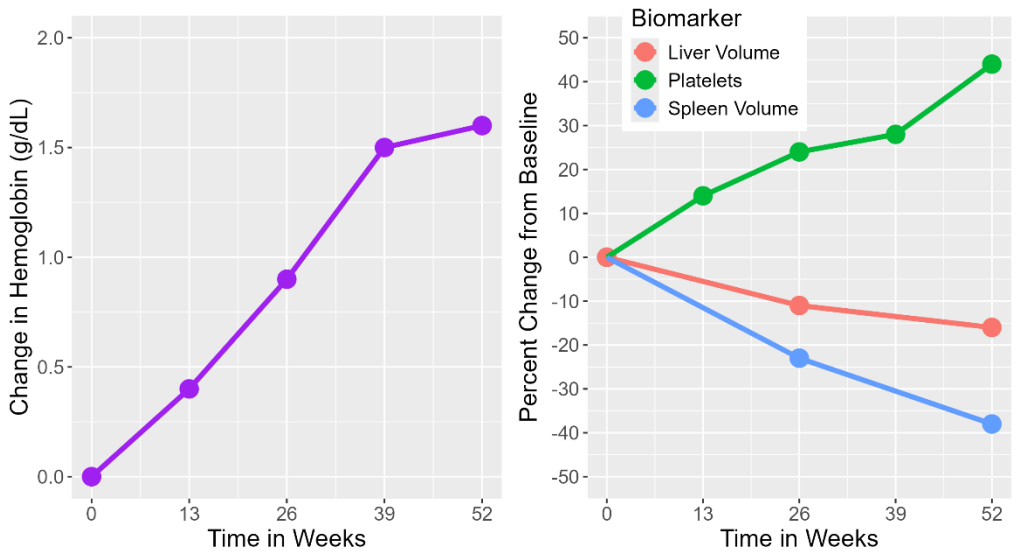


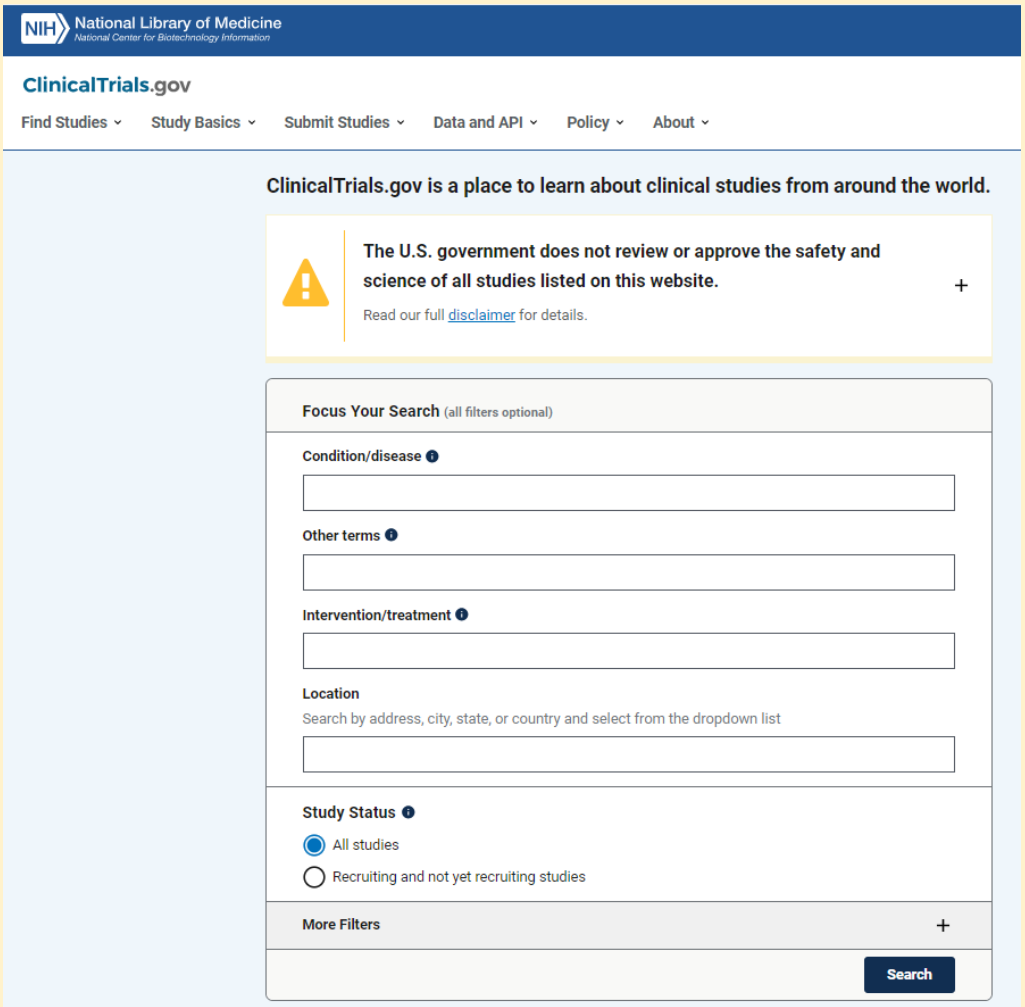
Figure 61. Biomarker changes following 50 or 100 mg twice-daily eliglustat administration in patients with Gaucher disease. All biomarkers indicated disease improvement. Figure modified from [200].

Returning to our eliglustat example, after completing the SAD, MAD, and food effect study, the company initiated a Phase 2 study in patients with Gaucher disease, an inherited metabolic disorder in which patients have a deficiency in the enzyme glucocerebrosidase that results in the accumulation of the lipid glucocerebroside in their liver and spleen. Patients with Gaucher disease have enlarged livers and spleens that lead to anemia, lung problems, bruising, and fatigue. Severe forms of the disease may lead to death. Currently, Gaucher disease is most commonly treated with intravenous enzyme replacement therapy, i.e., exogenous recombinant glucocerebrosidase administered every 2 weeks for life. The standard of care at the time was imiglucerase (Cerezyme), a product made by the same company. Eliglustat was to be an oral version of Cerezyme.

In their Phase 2 study, 26 adult Gaucher patients (16 females and 10 males) were enrolled; 22 completed the study [200]. Based on animal models, it was predicted that plasma concentrations of 6-14 ng/mL would be required for efficacy in patients. It was decided that 50 mg twice daily would achieve the desired blood concentrations. However, given the wide inter-patient variability in plasma concentrations, there was concern that some patients would not be able to achieve the target concentration. For those patients, a higher dose was needed. In the study, patients were administered 50 mg of eliglustat twice-daily for 52 weeks. On Day 20, patients had a blood sample collected and assayed for their plasma eliglustat concentration. Patients with concentrations less than 20 ng/mL had their dose increased to 100 mg twice-daily. Patients with Gaucher disease have an enlarged spleen and liver with low hematocrit and platelets. Changes in the parameters were assessed at specific times during the study. Safety was assessed throughout. Consistent improvements in hemoglobin, platelet counts, spleen volume, and liver volume were observed at each visit (Figure 61). The study was considered a success.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

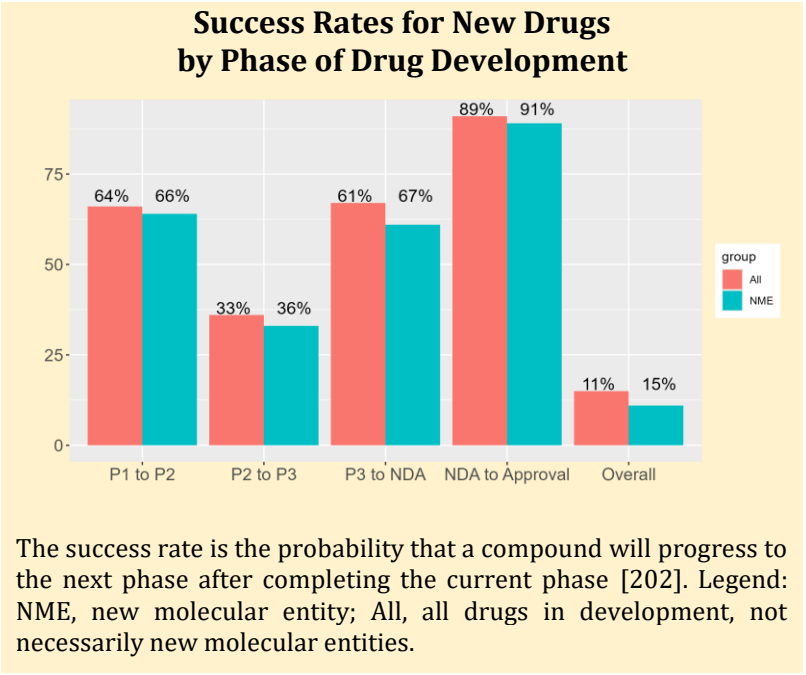
ClinicalTrials.gov - Your One-Stop Shop for Clinical Trial Information



ClinicalTrials.gov is a public website repository maintained by the US National Library of Medicine that houses information related to the design of ongoing and completed clinical trials in human volunteers, as well as some high-level study results. Developed as part of the FDA Modernization Act of 1997 and going live in 2000, its purpose was to provide patients with greater access to clinical trials and experimental therapies. Today, it houses information on nearly all clinical trials in the US and over 200 countries worldwide [201]. Although not all clinical trials are legally required to be posted on the website, nearly all studies are. As of March 2021, the repository housed almost half a million studies, with 30% located in the US and 54% originating outside the US. More than 60,000 of these studies were active and recruiting patients! Although many of the studies on clinicaltrials.gov are pharmaceutical studies, more than half are observational studies that involve no specific treatment intervention, behavioral studies, or surgical procedures or medical devices. In 2007, the law and site were expanded, requiring sponsors to post their clinical trial results and adverse events. As of 2025, however, only ~13% of trials have results posted on the site.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Statisticians are becoming quite clever at designing more efficient clinical trials. More on the nitty-gritty of clinical trials, their design, and their analysis can be found in Appendix 1: Design Considerations for Clinical Trials.



The success rate is the probability that a compound will progress to the next phase after completing the current phase [202]. Legend: NME, new molecular entity; All, all drugs in development, not necessarily new molecular entities.

Phase 3

Once the drug dose has been optimized in Phase 2, sponsors will move on to large-scale Phase 3 clinical trials in patients. These trials are RCTs, often double-blind, involving an appropriate control group. These studies are often international in scope, conducted across many clinics within a country, have rigorous inclusion/exclusion criteria, and may involve thousands of patients. For these studies, showing a change in a biomarker, as in Phase 2, is often insufficient for drug approval. Sponsors must demonstrate benefit on some clinically meaningful endpoint. For instance, in cancer, it is not enough to show that your drug makes tumors smaller; the sponsor has to show that the drug prolongs survival – you live longer taking the drug. Even with all the work done to this point, about one in three drugs fail Phase 3, usually because they do not show efficacy, but sometimes because an adverse event that was not previously observed in a larger population begins to occur.

Continuing with the eliglustat example, having demonstrated a positive Phase 2 outcome, the sponsor designed a head-to-head comparison to the gold standard of care at the time, imiglucerase [203]. In this study, adults who had received imiglucerase for 3 years or more were eligible for enrollment. Patients were stratified by imiglucerase dose (low vs high) and then randomized 2:1 to either their stable imiglucerase dose or eliglustat 50 mg twice daily. After 4 weeks of treatment, patients were assessed for eliglustat concentrations; those with low concentrations had their dose increased to 100 mg twice daily. The treatment duration was 1 year. If eliglustat works as planned, their hematological parameters and organ volumes should remain stable over time after transitioning from imiglucerase therapy to eliglustat treatment.

A total of 106 patients were randomized to receive eliglustat, and 54 to receive imiglucerase. The per-protocol population was analyzed, comprising randomized patients who demonstrated>80% compliance with their medication regimen, had no major protocol deviations, and showed no significant decline in hemoglobin or platelet counts due to disorders other than Gaucher's disease. A total of 85% of eliglustat patients and 94% of imiglucerase patients remained stable over the 12-month period (Figure 62). The difference between the rates was -8.8%, within the interval deemed by experts to be similar in effect to that of imiglucerase. Hence, it was concluded that eliglustat was equivalent to imiglucerase in maintaining stable values. The sponsor filed a New Drug Application for regulatory approval based on these results.

As with everything in life, there are exceptions to the development paradigm just described. Some drugs, like cancer drugs, cannot be given to healthy volunteers for safety reasons. Phases 1 and 2 may be combined for these drugs, and the SAD study may be skipped altogether. In the United

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

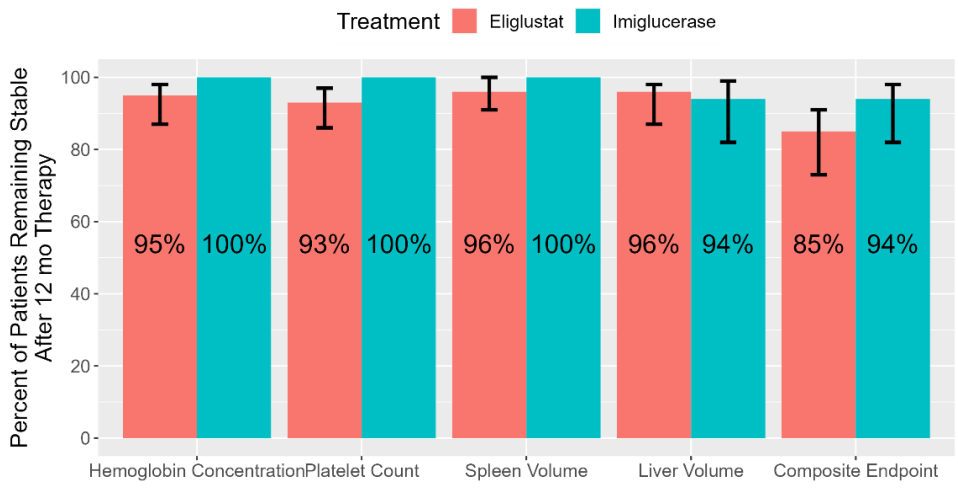


Figure 62. Percentage of Gaucher patients who remained stable in hematologic parameters and organ volumes following 12 months of treatment with imiglucerase or eliglustat. The composite endpoint was the percentage of patients whose hematologic parameters *and* organ volumes remained stable for 12 months. Figure redrawn from [203].

States, the Food and Drug Administration requires "at least two adequate and well-controlled clinical trials" for approval. For some rare diseases, it may be impossible to conduct adequately controlled trials because the population is too small. For instance, Myozyme (alglucosidase alfa) was approved based on two studies: an open-label study in 18 patients with Pompe disease and another in 21 patients. Hence, the drug was approved based on data from 39 patients.

However, there are only 5,000-10,000 Pompe patients worldwide, so 39 patients were quite a lot (approximately 1 in 200 patients with the disease enrolled). There are an estimated 500 million people in the world with Type II diabetes. Running a similar-sized study in patients with diabetes would require a clinical trial involving 26 million patients! Recent history suggests that, in the context of cancer and rare diseases, the two-controlled-trial rule can be broken. IQVIA recently reported that in 2018, nearly half of all new drug approvals relied on only one trial (29 of 52 new drugs); specifically, 1 in 8 relied on results from Phase 1 or 2 trials with no subsequent Phase 3 trials, with most of these being in cancer or rare disease indications.

Near the end of Phase 2b and the start of Phase 3, sponsors will revisit and complete many more Phase 1 studies to determine the optimal dose of the drug in different populations or when other drugs are coadministered. The following are the types of Phase 1 studies done to complete the registration package:

- Drug interaction studies. A comparison of the drug's pharmacokinetics in the presence and absence of another drug. There may be upwards of a dozen of these studies conducted.
- Mass balance study. A radioactive version of the drug is administered. The process by which the drug is metabolized and excreted into the blood, urine, feces, and expired air is determined.
- Food effect. An adequately powered study is designed to investigate the effect of a meal on a drug's pharmacokinetics.
- Renal impairment study: A study to determine the drug's pharmacokinetics in otherwise healthy volunteers who have varying degrees of renal impairment (mild, moderate, severe, end-stage, dialysis).
- Hepatic impairment study. It is the same as renal impairment, but with volunteers who have varying degrees of hepatic impairment.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

When Clinical Trials are Cancelled

I was watching an episode of Grey's Anatomy recently. In that episode, a patient with terminal cancer, who had failed all other prior therapies, was enrolled in a clinical trial for a new drug, and after a few weeks of treatment, her tumor shrank by 30%. Everyone was thrilled. But then, an email arrived notifying the hospital that the clinical trial was being cancelled due to a lack of funds. Although this is a fictional story, it is based on reality when, in 2025, the Trump administration cancelled almost 400 clinical trials affecting almost 75,000 people [204]. Although these numbers are shocking, clinical trials are cancelled for many reasons, and it has been estimated that about 1 in 10 clinical trials are cancelled while the study is ongoing. This raises the question, what happens to participants when their clinical trial is cancelled?

The answer depends on the reason for cancellation. If the trial was stopped for safety reasons, all patients may be prohibited from receiving further treatment, regardless of the outcome. This happened in 2006 to torcetrapib, a drug being developed to treat high cholesterol, in the 15,000-patient Illuminate trial. An increased risk of death and serious cardiac events was detected in the torcetrapib group, and the trial was terminated with all patients told to stop taking the drug immediately [205]. Stoppage can also occur with cancer drugs, even if the patients are responding to treatment. In 2014, idelalisib was approved for the treatment of certain blood cancers. However, in 2016, many clinical trials with idelalisib were stopped mainly due to an increased risk of infection and serious adverse events. In 2022, it was withdrawn from the market.

If the trial was stopped due to lack of efficacy, perhaps because an interim analysis showed that the response rate of the drug was too low for regulatory approval, then often companies will allow those patients who had responded to the drug to continue receiving the drug through a compassionate use or expanded access program.

If the trial was stopped for administrative reasons, such as the company has decided to reprioritize its portfolio or make a strategic decision not to continue development of the program because the financial considerations are such that it is not cost-effective or it would be unlikely the drug would get regulatory approval, then sponsors often allow access to the drug through compassionate use or expanded access, but for a fixed period of time. Such was the case with Sanofi, which terminated a Phase 3 trial with amcenestrant, a drug being developed to treat breast cancer, when a concurrent Phase 3 trial did not demonstrate any benefit, and the likelihood of drug approval decreased from 42% to 17% [206].

If the trial was stopped because the funding was unavailable, then patients will likely be unable to receive any drug in the future once supplies run out at a site, even if they were responding to the drug. Such was the case with The Immune Response Corporation (TIRC), a company co-founded with Jonas Salk, inventor of the Polio vaccine. In 1996, TIRC initiated Trial 806, a 2,527-patient Phase 3 trial in HIV patients that tested whether immune stimulation could slow the progression to AIDS or death. During the study, antiretroviral drugs were introduced, and suddenly, there was little interest in immune stimulation as a therapy for HIV. Having difficulty securing funding to complete the study, the study was terminated.

So what about all those patients enrolled in clinical trials terminated by the Trump administration in 2025? While some were able to restore funding through legal action, many patients were left in limbo, unable to obtain any further treatment, even if that treatment showed benefit in some patients.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

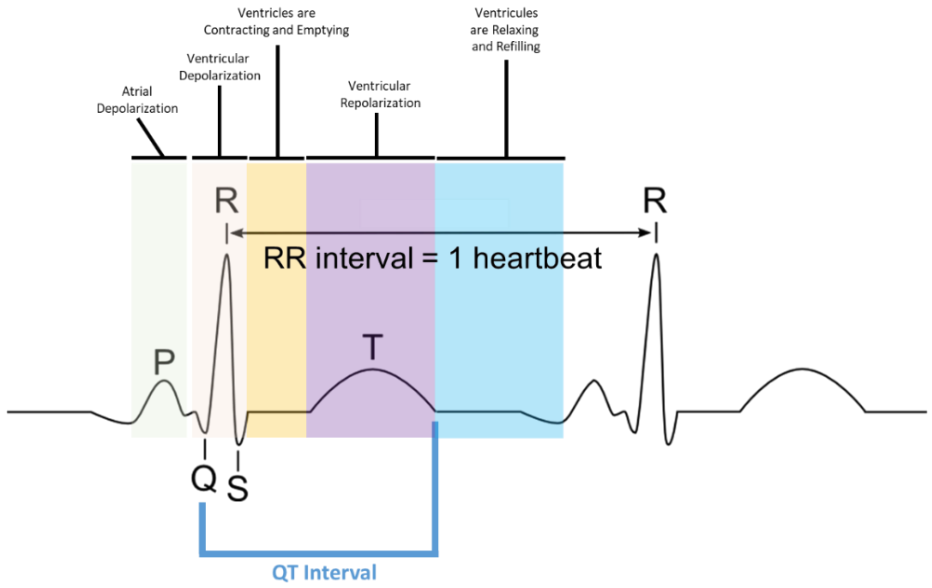


Figure 63. Electrical and mechanical events on an ECG. Some drugs, like Cisapride and Seldane, have a high risk of arrhythmias by blocking a K+-specific ion channel called HERG, preventing potassium ions from leaving cardiac cells and preventing the heart from properly repolarizing after a beat. Repolarization manifests as prolongation of the QT interval on an ECG. QT studies, in general, are designed to assess whether a drug prolongs the QT interval and may be at increased risk of arrhythmic liability.

- Thorough QT trial. This study is designed to assess the drug's effect on heart repolarization by examining whether it prolongs the QT interval on an ECG (Figure 63). Several drugs were removed from the market (Seldane, Cisapride) because they sometimes led to a fatal arrhythmia called Torsades de Pointes (TdP). These studies are designed to weed out drugs with a high potential for TdP.
- Formulation studies. A comparative study was designed to assess the bioequivalence of a new formulation compared to another formulation.

Returning to eliglustat, once positive Phase 2 results were obtained, and while Phase 3 was being conducted, the sponsor conducted a flurry of Phase 1 studies:

- A bioequivalence study in healthy volunteers to compare the pharmacokinetics of the final to-be-marketed formulation versus the formulation used in clinical trials (NCT01452542);
- A drug interaction study in healthy volunteers with digoxin and metoprolol (NCT01357811 and NCT01659944);
- A pharmacokinetic study in in healthy volunteers who are also CYP2D6 poor metabolizers compared to healthy volunteers who are also CYP2D6 extensive metabolizers (NCT06188325);
- A taste evaluation study in healthy volunteers of different liquid formulations of eliglustat (NCT02422654);
- A pharmacokinetic study in healthy volunteers comparing the absorption of an oral tablet to an oral solution (NCT06193304);
- A pharmacokinetic study in volunteers with renal impairment (NCT02536937);
- A pharmacokinetic study in volunteers with hepatic impairment (NCT02536911); and
- A mass balance study in healthy volunteers (NCT06143904).

These studies rounded out the clinical pharmacology section of the eliglustat package insert.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Modeling and Simulation in Drug Development

Mathematical modeling and simulation play a pivotal role in drug development. The costs of developing a new drug continue to escalate, and companies must seek ways to reduce these expenses. One particularly useful tool is mathematical modeling. Scientists in drug companies, known as pharmacometricians, will develop mathematical models that relate drug concentrations in the body to a measure of the drug's effect. Using simulation, they will optimize the dose to maximize efficacy while minimizing adverse events. When properly validated, these models can be used as an alternative to conducting a clinical trial, thereby saving a company millions of dollars.

An example of this is the approval of paliperidone palmitate for the treatment of schizophrenia. A long-acting, extended-release tablet was already approved, and the sponsor approached the FDA to seek approval for an injectable once-monthly formulation. Mathematical models were used to recommend the dosing regimen, dosing window, strategy for handling missing doses, method for switching from prior treatments, dosing approach for patients with concomitant renal impairment, and dosing regimen for other special populations. No clinical studies were conducted! All of these recommendations are in the product label.

movantik® (naloxegol) 25 mg tablets

Another example is the use of physiologically based pharmacokinetic (PBPK) models, which describe drug distribution in the body using anatomically and physiologically accurate mechanisms. Using PBPK models, sponsors may not have to do certain drug interaction studies. For example, Movantik (naloxegol), a drug used for the treatment of opioid-induced constipation, is co-administered with rifampin, a drug that increases the expression of drug-metabolizing enzymes and increases the metabolism of many drugs, leading to reduced drug concentrations and reduced efficacy. They found significantly reduced naloxegol concentrations in the presence of rifampin, such that coadministration of rifampin is not recommended with naloxegol. Rifampin is classified as a strong inducer due to its potent effect. The question then becomes, what about moderate inducers, such as efavirenz (Sustiva)? How should those be dosed? Using PBPK modeling, the sponsor simulated a study involving the coadministration of 25 mg naloxegol with efavirenz and found that the simulated blood concentrations were similar to those observed when naloxegol was administered at the lower dose of 12.5 mg. Based on this information, the physician may prescribe a higher dose if the drug does not appear to be working. There are many other examples of modeling and simulation to improve decision-making and help prescribers.

Post-Marketing Requirements and Phase 4 Studies

Studying the drug does not stop when the drug is approved for marketing. The regulatory agency may return and require additional studies as a condition for approval (so-called post-marketing commitments or requirements). Collins et al. [207] examined all 132 novel oncology drugs approved by the FDA between 2012 and 2022 and found that the FDA required 376 post-marketing commitments. Most of these commitments required the company to study the effects of patient-related factors (so-called intrinsic factors), such as weight and genetic factors, and external factors (so-called extrinsic factors), including food and drug interactions, on dosing. Companies were given a median of 2 years to complete these studies. In some cases, studies were required to examine the efficacy of lower doses or the safety and efficacy of alternative dosing regimens. In some cases, additional safety studies, such as examining the effect of the drug on QT intervals on an ECG, were required, as were additional animal toxicology studies.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Once the drug is approved, the company may begin studying the drug in other patient populations to expand its label. A company may discover that a drug interaction it never considered during development appears to occur based on anecdotal evidence from adverse event logs. In such cases, the company may conduct a specific drug interaction study to determine whether this is indeed the case and whether the dose needs to be adjusted when co-administered. If the drug has been approved for a long time, the company may develop more friendly, alternative formulations. If a drug has to be taken twice a day, the company might develop a once-a-day formulation. These are so-called lifecycle management studies designed to prolong the drug's profitability as it nears the end of its patent period. These types of studies are sometimes called Phase 4 studies.

One particular post-marketing study warrants an in-depth discussion, specifically identifying the appropriate dose in children. Under the Best Practices in Children Act of 2002, drug companies have been encouraged to find the optimal dose in children and, in return, will be granted an additional 6 months of patent exclusivity. Sponsors may waive out of pediatric studies if children will not be administered the drug. For example, fezolinetant is used to treat hot flashes in women with menopause, a condition that typically affects women in their 50s. The sponsor opted out of pediatric studies because the likelihood of children being prescribed the drug is virtually zero.

How do you find the best dose in children? There are old-time rules of thumb, like Young's Rule:

$$\text{Pediatric Dose} = \left(\frac{\text{age of child in yrs}}{\text{age of child in yrs} + 12} \right) (\text{adult dose}) \tag{32}$$

Or Clark's Rule:

$$\begin{aligned} \text{Pediatric Dose} &= \left(\frac{\text{weight of child in lbs}}{150 \text{ lbs}} \right) (\text{adult dose}) \\ &= \left(\frac{\text{weight of child in kg}}{68 \text{ kg}} \right) (\text{adult dose}) \end{aligned} \tag{33}$$

If the adult dose were 100 mg for a 12-year-old weighing 90 lbs., the pediatric dose would be 60 mg under Clark's rule and 50 mg under Young's rule. This dose might then be rounded up or down if that strength dose was not commercially available. However, these are just guidelines and not particularly precise.

Today, a more rational approach is used. The first step is to determine whether the illness in children has the same pathophysiology as in adults and whether children will respond to treatment in the same way as adults. If the answer is 'no,' extensive dose-finding studies are required, similar to those conducted in Phase 1 and 2 drug development in adults. If the answer is 'yes,' exposure matching can be used. The idea behind exposure matching is that if you know the therapeutic dose in adults and the corresponding plasma drug concentrations at that dose in adults, then first estimate a dose in children that produces the same concentrations as in adults, and then do a clinical trial in kids to see if the predicted dose does produce equivalent exposure as adults.

If the observed concentration data in kids match those in adults, that is the recommended pediatric dose. If the exposure does not match, then the observed data are used to make a more accurate prediction of the pediatric dose.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

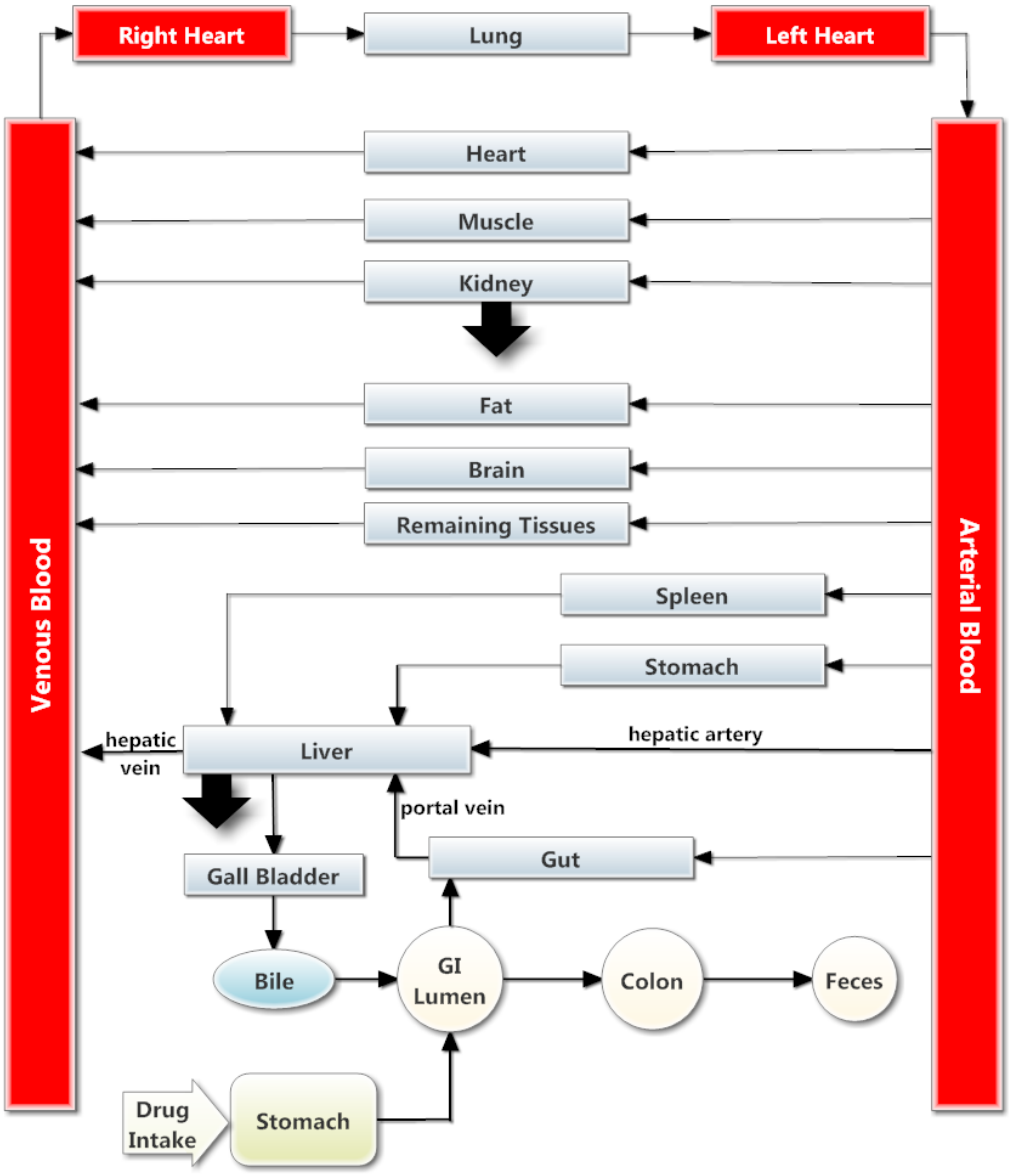


Figure 64. Example of a physiologically-based pharmacokinetic (PBPK) model. PBPK models treat the body as a series of interconnected compartments (organs), where the rate into and out of each compartment is determined by blood flow. The volume of distribution for each compartment is based on the organ's actual volume. These models are useful because they account for immature physiological processes, such as enzyme maturation. A model for children can be estimated by developing a model for adults and scaling down the organ volumes and blood flows to the size of children, while also accounting for any immature physiological processes that may occur at the age of interest. Drug concentrations can be simulated using a hypothetical dose. The dose in children is the one that produces concentrations equivalent to those in adults at the therapeutic dose.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

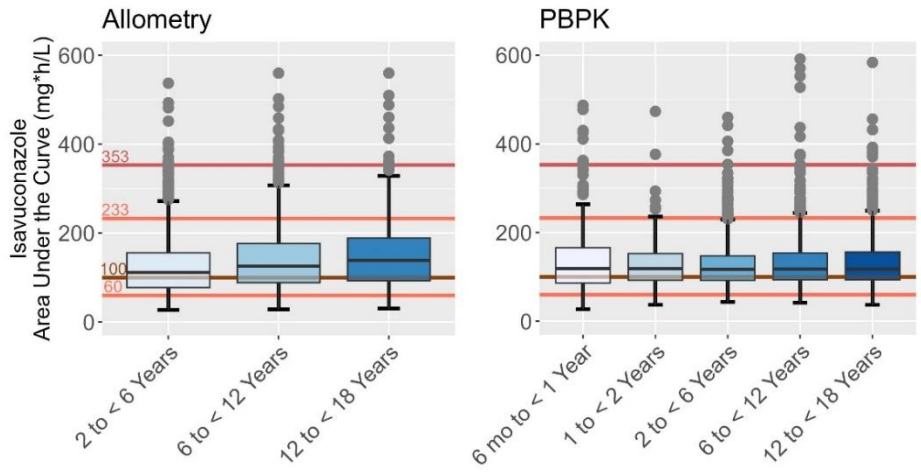


Figure 65. Predicted isavuconazole AUC in children of different ages administered isavuconazole sulfate. A total of 1000 children were simulated in each age group. Children 2 years or younger were administered 6 mg/kg, while children older than 2 years were administered 10 mg/kg with a maximal dose of 372 mg. The boxes contain 75% of the simulated children, and the whiskers approximately span the 5th and 95th percentiles of the data. The line inside the box is the mean value. The solid reference lines indicate the target levels. Figure modified from [208].

Two different methods are generally used to find the pediatric dose:

1. The first method utilizes allometric scaling, whereby the drug's pharmacokinetics scale with body weight. By scaling clearance and volume of distribution and varying the dose, thousands of children of different ages and weights can be simulated, generating a range of concentrations at a particular dose. The dose that most closely matches the exposure seen in adults is then used as the basis for the clinical trial. This approach works well down to 2 years of age, but below that, it does not work as well because many physiological functions in children have not yet reached adult levels. For example, CYP3A4, the enzyme responsible for the metabolism of most small-molecule drugs, does not fully mature until about 2 years of age. Hence, immature physiology must also be taken into account when predicting pediatric doses.
2. In children less than 2 years of age, physiologically-based pharmacokinetic (PBPK) models are used (Figure 64). PBPK models treat the body as a circulatory model with blood flowing into and out of the organs. Organ sizes refer to their actual anatomic volumes. Elimination is accounted for by hepatic and renal clearance, and any immature physiological processes can also be modeled. Using the PBPK model, children of a given weight and age are simulated, and the corresponding drug concentrations are simulated at a given dose. The pediatric dose is the dose that results in drug concentrations equivalent to those administered to adults at the therapeutic dose.

As an example, Cresemba (isavuconazonium sulfate) was approved in 2015 for the treatment of aspergillosis, a rare but serious fungal infection. Isavuconazonium sulfate is a prodrug – it has no pharmacological activity on its own, but is metabolized to an active metabolite called isavuconazole, which does have pharmacological activity. The metabolism of isavuconazonium sulfate is so fast and complete that it is undetectable in blood; the entire absorbed dose is metabolized to isavuconazole.

To estimate a safe and suitable dose of isavuconazole in pediatric patients, it was first determined that the pathophysiology of aspergillosis is similar in adults and children. Hence, an exposure-matching approach could be taken. Two methods for pediatric dose estimation were used: allometry

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 10 Recommended Dose of Cresemba in Pediatric Patients			
Age (yrs)	**Body Weight (kg)**	**Loading Dose**	**Daily Dose Thereafter**
1 to < 3 years	< 18 kg	15 mg/kg IV every 8 hours for 48 hours (6 doses)	15 mg/kg IV once-daily
> 3 years	< 37 kg	10 mg/kg IV	10 mg/kg IV once-daily
The maximal dose is not to exceed 372 mg.			

for 2-18 years of age and PBPK for 6 months-18 years of age (Figure 65) [208]. Allometry and PBPK models were developed in adults and scaled down to children. To ensure the correlation between age and weight was maintained, a total of 1000 children were randomly sampled for each age group from the National Health and Nutrition Examination Survey (NHANES), a database of health parameters like age, weight, height, laboratory values, etc., for thousands of people across the United States and curated by the Centers for Disease Control. The dose administered was varied by trial and error to find the best value subject to the following targets and constraints:

- The mean simulated value in children should be close to the mean target AUC value of 100 mg*h/mL in adults.
- The majority of values in children should be within 60 and 233 mg*h/mL, which were the smallest and largest observed values in adults administered the therapeutic dose.
- Ideally, no patient should exceed 353 mg*h/mL, which was the lowest AUC where adverse events started to occur in adults when administered higher than therapeutic doses.

Under these conditions, the best dose was estimated to be 10 mg/kg in children 2 years and older and 6 mg/kg in children younger than 2 years. Based on these predictions, a clinical trial in children was initiated (NCT03241550) in 2017 and completed in 2019. Children with aspergillosis were administered Cresemba orally or intravenously. Blood samples were collected and analyzed for isavuconazole concentrations. The AUC was then estimated for each child and compared to the target values and constraints. Based on these results, the pediatric dosing regimen was decided and is shown in Table 10. The 10 mg/kg dose was predicted quite well, but the pediatric dose in children < 3 years of age needed to be increased because the observed AUC in these children was lower than predicted. This label change was approved by the FDA in 2024.

Lifecycle Management of Drugs

Patents have a minimum 20-year life. After considering how long it takes to develop the drug and bring it to market, that often doesn't leave manufacturers much time to recoup their costs. After the patent period has expired, generic manufacturers are free to make generic copies of the product. Finding and developing new drugs to replace those lost to patent expiration is challenging and costly. It's far simpler and cheaper to find ways to prolong sales of the already established product. This has led companies to develop a range of lifecycle management strategies to sustain product sales. Five strategies are particularly prominent:

- Reformulations,
- Chiral switches,
- Combination products, and
- Over-the-counter switches.
- Legal maneuvers

Each will now be discussed.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

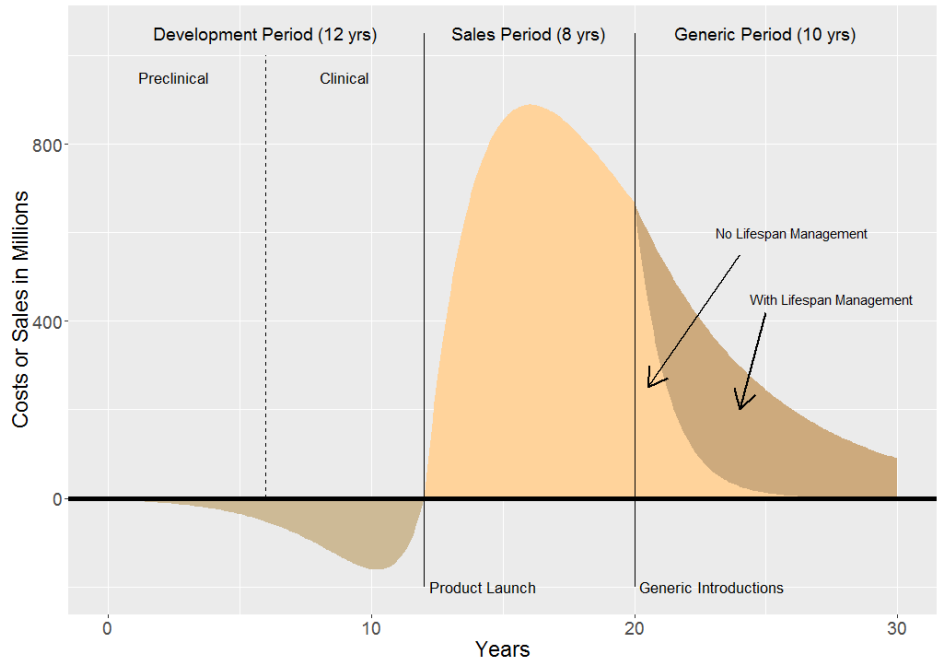


Figure 66. Cost (below the 0 line) or sales (above the 0 line) of a first-in-class blockbuster drug. Once the generic period is entered, lifecycle management can prevent sales losses from generics. This profile's shape has been called a “shark-fin”.

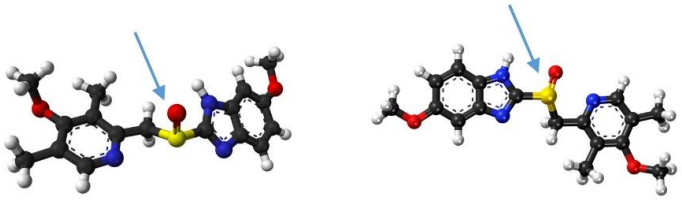
Reformulation. Most small-molecule drugs are administered orally using an immediate-release (IR) formulation, which means that as soon as the formulation is in the stomach and comes into contact with a solution, such as stomach contents, it disintegrates. A drug cannot be absorbed into the body until it is dissolved, which means the formulation must first disintegrate. One lifecycle management strategy is to modify the IR formulation to leverage drug-delivery technologies and feature a different release rate. These time-release formulations may be absorbed faster, such as with oral dispersion tablets that are absorbed buccally or sublingually, or may be designed to disintegrate and be absorbed more slowly, such as with extended-release, controlled-release, or sustained-release formulations, which are all synonyms for the same concept. These extended-release formulations may enable manufacturers to switch from twice-daily dosing to once-daily dosing for short half-life drugs. For example, Pfizer’s drug Xeljanz (tofacitinib citrate), used to treat rheumatoid arthritis, is an IR tablet dosed twice daily at 5 mg. Pfizer reformulated Xeljanz as an 11 mg modified-release tablet, now marketed as Xeljanz XR. Another example is Boniva (ibandronate sodium), which is used to treat osteoarthritis. Boniva was originally marketed in 2003 as a 2.5 mg intravenous (IV) tablet to be administered once daily. But in 2008, Roche reformulated Boniva into a 150 mg once-monthly tablet. Other formulation changes could involve a complete switch in route of administration, such as moving from oral to a transdermal patch or a nasal spray.

Chiral Switches. In 1989, AstraZeneca introduced Prilosec (omeprazole) as a novel treatment for heartburn. Unlike H2-histamine antagonists, which had been the mainstay of therapy for years, omeprazole had a novel mechanism of action: it was an irreversible inhibitor of H2-receptors and was effective at blocking gastric acid secretion. Prilosec quickly became one of the best-selling prescription medicines in history, earning the company billions of dollars in revenue. But in 2001, the patent for omeprazole was set to expire. Yet when that date was reached, no generics were launched. Why? Because of the company's careful planning.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The "Shark Fin Project," named after the shape of the sales curve (Figure 66), was a strategic initiative to protect its blockbuster drug. AstraZeneca employees identified almost 50 possible solutions to prevent this from happening. The best idea was to find a better drug to replace Prilosec with a drug of similar efficacy but with more years of patent life left, a so-called "me-too" drug. The latter solution was not ideal because it would then compete against a cheaper generic drug with similar efficacy, which would severely limit sales [209].

Structure of Omeprazole and Esomeprazole



Omeprazole Esomeprazole

Each ball is an atom. Black balls are carbon, blue is nitrogen, yellow is sulfur, red is oxygen, and white is hydrogen. The arrow points to the location of the chiral center, which is the sulfur atom.

Omeprazole is a chiral drug. It exists as a 1:1 mixture of two enantiomers, designated R- and S-. AstraZeneca scientists then began examining the enantiomers. In vitro studies failed to show any difference between the enantiomers, and animal studies were inconclusive: rats showed a different effect with one enantiomer, but dogs did not. At this point, AstraZeneca rolled the dice and initiated human studies. Normally, a company would probably not invest the money because the in vitro and animal results were so underwhelming, but with so much on the line, the decision was made. In humans, differences were observed between the enantiomers. S-omeprazole, which was later renamed esomeprazole, had higher plasma concentrations than omeprazole at the same dose. When gastric acid secretion was stimulated in healthy volunteers with the drug pentagastrin, esomeprazole showed 91% inhibition, compared with 65% for omeprazole and 25% for R-omeprazole, after oral administration of 15 mg. In patients with gastroesophageal reflux disease (GERD), esomeprazole demonstrated higher healing rates and greater symptomatic relief than omeprazole. This led AstraZeneca to seek approval for esomeprazole, which was granted in 2001 and launched as Nexium, becoming a blockbuster drug. The launch of Nexium prevented the introduction of generic omeprazole because doctors would be unlikely to prescribe an inferior product if a better one were available.

This story is typical of chiral switches, in which a single enantiomer is obtained from a racemic mixture. In contrast to some other lifecycle management options, a chiral switch involves developing a completely new drug in the eyes of the FDA. And rightfully so, since it has long been known that enantiomers have different chemical and pharmacological properties. The difference is that the approval risk is lower, since you already know the safety and efficacy of the racemic product. Other examples of racemic switches are levofloxacin, levalbuterol, and escitalopram.

Fixed Dose Combinations. Another way manufacturers have found to prolong the sales life of their products is through the development of fixed-dose drug combinations, where two or more drugs are formulated into a single tablet. This approach aims to improve patient compliance by reducing the total pill burden and increasing ease of use. Examples include:

- Antihypertensive combinations like Prestalia, which is the combination of the Angiotensin Converting Enzyme (ACE) inhibitor perindopril and the calcium channel blocker amlodipine.
- Antidiabetic combinations like Glucovance, which is the combination of the sulfonylurea glibenclamide and metformin.
- Dual antiretroviral combinations like Prezcobix, which is a combination of the protease inhibitor darunavir and the cytochrome p450-3A inhibitor cobicistat, and triple combinations like Atripla, which consists of the non-nucleoside reverse transcriptase inhibitor (RTI)

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

efavirenz, the nucleoside RTI emtricitabine, and the nucleotide analog RTI tenofovir.

It is worth noting that, in the FDA's view, fixed-dose combinations differ from combination products. Combination products are technically any combination of two or more regulated products, like a drug and a device. What is interesting about this definition is that the combination of a drug and a biologic is considered a combination product, not a fixed-dose combination product. Fixed-dose combinations refer to two products of the same type.

Some Fixed-Dose Combination Products Work Better Than Their Single Agent Components

Intraocular Pressure Reduction After 3 Months Treatment In Patients With Glaucoma or Ocular Hypertension			
Time of Day	Fixed-Dose Combination Simbrinza®	Single-Agent Brinzolamide	Single-Agent Brimonidine
8 am	-26.6%	-22.6%	-17.1%
10 am	-34.9%	-22.4%	-25.8%
3 pm	-24.1%	-16.9%	-14.3%
5 pm	-29.7%	-17.7%	-23.9%

Simbrinza is a fixed-dose combination suspension of 1% brinzolamide and 0.2% brimonidine. Reprinted from [210, 211].

Over-the-counter drugs (OTC). An OTC drug is a prescription medicine that is now sold directly to consumers without a prescription. Drug manufacturers may petition to have their drug changed to OTC. The reader is referred to the section on Over-the-Counter Drugs in the Regulatory Considerations chapter for further details and information.

Legal Maneuvers. Some companies resort to legal maneuvers to prevent generic competition from entering the market. For example, Namenda (memantine) was an IR formulation approved in 2003 for the treatment of Alzheimer's disease. With its patent set to expire in 2015, Forest Labs launched an extended-release (ER) version of memantine in 2013. Rather than have both the IR and ER versions on the market, Forest Labs pulled the IR formulation from the market in 2014, forcing memantine users to switch to the ER formulation. Patients rarely change formulations, and once a patient is switched to the ER formulation, they are unlikely to switch to a generic IR formulation. This maneuver prompted the New York State Attorney General to file a complaint against Forest Labs, preventing them from removing the IR formulation from the market [212]. The courts agreed and issued an injunction against Forest Labs, preventing them from removing the IR formulation from the market until August 2015, 1 month after generic formulations became available.

Investigator-Initiated Trials

Drug companies are not the only ones that study a drug. Once a drug is released and available to the public, anyone can use the drug for research. Many academic scientists study the effects of drugs in various real-world settings under what are called Investigator-Initiated Studies or Trials (IITs). The funding for these trials can come from federal grants, the academic institution where the investigator works, or the pharmaceutical company that markets the drug. Many pharmaceutical companies have standard operating procedures (SOPs) that govern when, how much, and under what conditions they will contribute to an IIT. Sometimes, the company provides the drug; other times, it provides a grant to cover the research costs. IITs provide additional information on drugs because drug companies cannot conduct every potentially relevant study. An example of an IIT is a physician who wishes to study the effect of a cancer drug on patients with a cancer not approved in the label.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 11 Three Examples of Registry Studies Randomly Selected From Clintrials.gov	
Study Name	**Objective**
International Collaborative Gaucher Group (ICGG) Gaucher Disease Registry & Pregnancy Sub-registry (NCT00358943)	• To enhance understanding of the variability, progression, identification, and natural history of Gaucher disease, with the ultimate goal of better guiding and assessing therapeutic intervention. • To assist the Gaucher medical community with the development of recommendations for monitoring patients, and to provide reports on patient outcomes, to optimize patient care. • To characterize the Gaucher disease population. • To evaluate the long-term effectiveness of imiglucerase and of eliglustat.
Assessing Barriers and Increasing Use of Immunization Registries in Pharmacies (NCT03796585)	• Identify barriers and best practices of immunization registry implementation, • Use a participatory design approach to develop an immunization registry training program, and • Disseminate and assess the impact of the immunization registry training program among community pharmacies' registry participation rates.
The Epidemiology of Patients Treated With ATZ+BEV in Real Life Setting (MANIFEST-HCC) (NCT06903663)	• To further investigate the clinical and demographic profile of patients receiving first-line treatment with atezolizumab and bevacizumab in the real-world setting of clinical practice from Romania

Registries

Patient registries are observational studies that collect uniform data from specific populations to evaluate specific objectives; they are considered to be real-world data. They may be used to describe the natural history of a disease, determine the effectiveness of therapeutic interventions, assess the safety of a therapeutic intervention, or evaluate the quality of life [213]. Examples of registry studies are found in Table 11. What a registry is can vary widely, depending on its objectives, and may be formed by investigators or at the request of regulatory agencies. Registries may be used as

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

supporting data for New Drug Applications or may stand alone for label expansions. The FDA has recently issued guidance for the use of registries to support regulatory decision making [214]. As with clinical trials, regulatory bodies look at how defined the population is, its inclusion/exclusion criteria, what data are collected, whether analytical methods are properly validated, how reliable the data, are and how relevant the data are to the proposed use.

An example of using a registry to support approval of drug is Spinraza (nusinersen). Spinal muscle atrophy (SMA) is a genetic disorder leading to muscle weakness, respiratory distress, and premature death; SMA patients rarely live longer than a few years. SMA is a rare disease and only about 8 children in 100,000 will have the disease. The Spinal Muscular Atrophy (SMA) Biomarkers in the Immediate Postnatal Period of Development (NCT01736553) was a registry study started in 2012 [215]. It's objective was to validate biomarkers of SMA and to develop a natural history of SMA patients from the first week of life to 2 years of age. By 2017, a total of 26 SMA patients and 27 control infants < 6 months old were enrolled in the registry.

Nusinersen is an antisense oligonucleotide that acts on the DNA to increase Survival Motor Neuron (SMN) protein levels, which are deficient in SMA patients. Because of the difficulty in identifying and recruiting SMA patients, Biogen used the registry study results as a natural history comparison in its Phase 2 study. Biogen asserted that the results indicated a survival benefit and clinical improvement compared to the natural history data. The FDA review was more cautious saying that confounding factors, such as differences in patient selection, made the data difficult to interpret. Nevertheless, Spinraza was approved in 2016.

Why Do Drugs Fail?

By now, it should be clear that not every drug tested in humans is approved for use. Most drugs studied in humans are never approved for use in humans, and they fail at some point in drug development. The drug is then terminated and ceases further development. But what are the reasons for failure? Before the 2000s, the major reason for drug failure was poor pharmacokinetics, such as the drug was not absorbed well, had a poor half-life, had erratic pharmacokinetics, etc. Through preclinical experiments and computer modeling and simulation, drug companies have largely eliminated poor pharmacokinetics as a reason for failure. In an analysis of 174 drugs that had failed in Phase 2 or Phase 3 between 2013 and 2015, nearly half failed for lack of efficacy, another quarter failed for poor safety, and the remaining failed for various other reasons:

- commercial considerations, such as the market has changed, such that the drug might not fit into the competitive landscape;
- strategic considerations, such as when two companies have merged and it no longer makes sense to develop the drug, or the company has shifted therapeutic areas; and
- operational reasons, such as it is too difficult to develop the drug or complete clinical trials in the proposed indication [216].

Looking back at earlier periods for other drugs, the authors concluded that a lack of efficacy remains the major reason for drug failure.

Even if a drug is not terminated, it may not be approved by regulatory agencies. Although not a rejection, when the FDA determines that a drug is not ready for approval in its current form, companies receive a Complete Response Letter (CRL). These letters are specific to deficiencies and may include recommendations to rectify any issues. These letters are intended to inform the sponsor of any issues that must be addressed before approval. Reasons may include questions surrounding efficacy, safety, manufacturing, or particular studies. The reasons do not necessarily have to be solely related to the sponsor; they may also involve third parties, such as third-party manufacturers. Sponsors may address the comments and resubmit the application.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

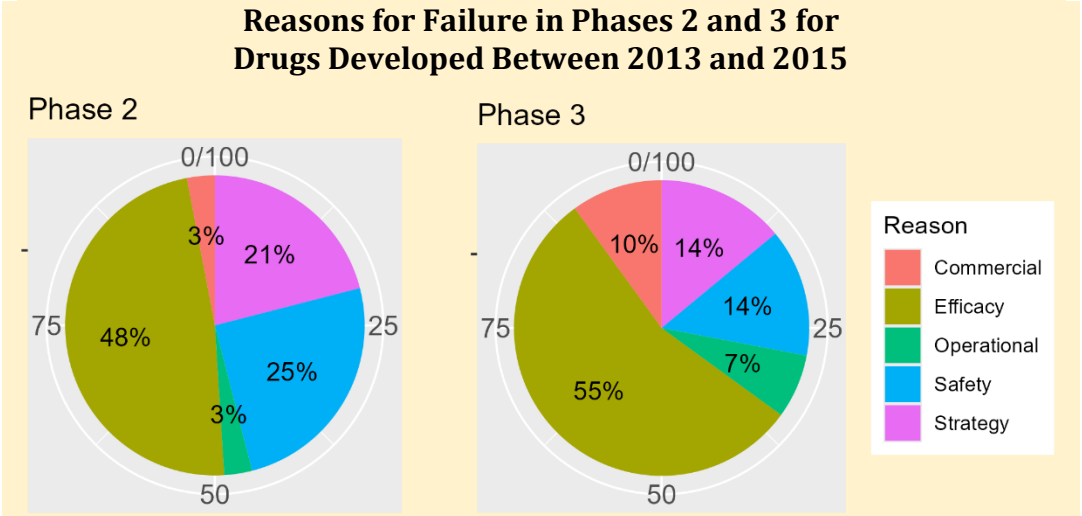


Figure redrawn from [216].

An analysis of CRLs from 302 New Drug Applications filed between 2000 and 2012 showed that only half of the initially submitted applications to the FDA were approved; however, the acceptance rate increased to 75% upon resubmission [217]. The reasons for rejection, as reported in this review, include questions about the proposed dose, the choice of study endpoints that failed to reflect a clinically meaningful endpoint, inconsistent results when different endpoints were compared, inconsistent results across studies or study sites, and poor efficacy compared to the standard of care.

Interestingly, in 2015, an analysis was conducted comparing 61 CRLs issued between 2008 and 2013 with subsequent press releases from companies whose drugs were not approved [218]. The authors concluded that the drugs failed for many reasons, most often due to deficiencies in both safety and/or efficacy. In most cases, companies did not issue a press release regarding their CRL; when they did, most statements about the CRL rationale were omitted. Additionally, of the press releases issued, 21% of the statements did not align with any information from the CRL. Only 14% of press release statements were considered accurate.

In 2025, the FDA published more than 200 CRLs for drugs submitted between 2020 and 2024, making them available to the public to improve transparency. A press release from the FDA cites [218] as the reason behind their issuance. Most of these were for already approved drugs and were already made public over the years, but they still provide an interesting peek at some of the reasons for rejection. A review by Fierce Biotech, an industry newsletter, found that some rejections stemmed from the usual reasons, including efficacy and safety concerns, third-party manufacturing issues, and inadequate data presentation. However, some of the reasons were a bit bizarre [219]. For example, Legubeti was rejected as a treatment for acetaminophen overdose by the FDA in 2023 because it had "an unpleasant odor that may affect the tolerability of oral ingestion." The FDA rejected the application because it was not convinced patients would take the entire dose at once due to the product's "bad" smell. Despite that, Legubeti was approved the next year.

Note: Most of the statistics in this section are relatively old, with some dating back more than a decade. No more recent analyses have been published, so it's unclear if these statistics and reasons still hold. I guess that these statistics still largely hold today.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Summary

Drug development is not a rigidly fixed process, as it might appear to be, as I have laid it out here. Nor is it a linear process. Drug development is so diverse that almost every drug has its own development plan. Setbacks must be overcome. In hindsight, development plans may appear straightforward and logical, but often they are not. One of the exciting aspects of drug development is the process of understanding the "personality" of your drug.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 7

Efficacy and Safety

Efficacy

With the passage of the FFDCA in 1938, sponsors were only required to demonstrate that their drug was safe. Following the thalidomide disaster in the early 1960s, the Kefauver-Harris amendment to the FFDCA required that a drug demonstrate both safety and effectiveness for approval. Evidence for effectiveness would now be based on "substantial evidence to support the claim" using "adequate and well-controlled studies" [2], which the FDA generally interpreted to mean two or more studies, each convincing in its own right and unlikely to result from chance alone. These guidelines remained in place for more than 30 years until the passage of the FDA Modernization Act of 1997 (FDAMA), where "one adequate and well-controlled clinical investigation and confirmatory evidence" was added to the definition of substantial evidence [220]. This section will discuss pivotal clinical trials and what makes them "adequate and well-controlled", how data from these trials are analyzed, and how they are reported. Real-world evidence, a more recent means to demonstrate safety and efficacy, will also be discussed.

What is an Adequate and Well-Controlled Trial?

The FFDCA (21 CFR 314.126) defines an "adequate and well-controlled study" [2]. Such studies are often called pivotal or confirmatory trials, because approval is pivotal to their outcome. Briefly, an adequate and well-controlled trial is one where:

- The objectives of the trial are clearly stated, as are the study design, duration of treatment, methods, what will be measured and assessed, and methods of analysis. This step occurs prior to any data analysis.
- The study design must allow a "valid comparison with a control to provide a quantitative assessment of the drug's effect." What constitutes a 'control' may be a concurrent placebo, different drug doses, active treatment by another drug, historical controls, or no control when the placebo effect is negligible.
- A sufficient number of patients having a predefined illness must be randomized to treatment groups such that bias is minimized with regard to the subjects, observers, or analysts of the data. For example, patients, observers, and analysts are all blinded to treatment.
- The methods of assessment must be well-defined and reliable.
- The trial is not poorly executed. For example, a high volunteer dropout rate could invalidate a study's results, as could premature unblinding by the analyst.

If the study is appropriately designed, the drug's effect can be distinguished from confounding factors, such as a placebo effect, and the study's results can be considered real, reproducible, and rigorous. Further expansion on what constitutes an adequate and well-controlled clinical trial can be found in a recent Guidance issued by the FDA [221].

Since FDAMA in 1997, the FDA has allowed one adequate and well-controlled trial when it is a large,

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026.

multicenter study supported with *confirmatory evidence* of effectiveness. These trials generally involve thousands of patients, are global in nature, and involve dozens of clinical trial sites, which can be hospitals or doctors' offices, each of which can have a significant impact on the outcome. The results need to demonstrate a "*clinically meaningful and statistically very persuasive effect on mortality, severe or irreversible morbidity, or prevention of a disease with potentially serious outcome, and with other characteristics described below, and confirmation of the result in a second trial would be impracticable or unethical*." With FDAMA, what constituted "confirmatory evidence" was not defined and was left to the FDA to decide what that meant.

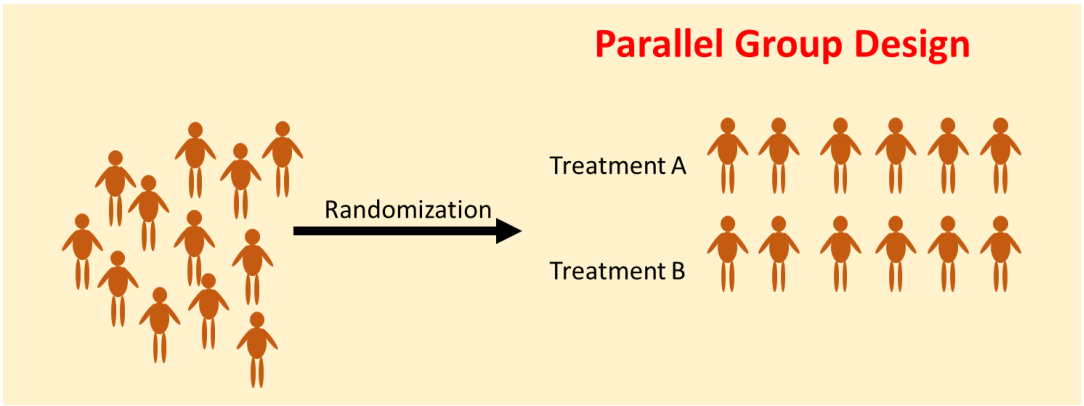
Although the FDA issued guidance on its interpretation of FDAMA in 1998 [222], what constituted confirmatory evidence was left largely undefined until later guidances were issued in 2019 and 2023 [221, 223]. Confirmatory evidence today in 2025 constitutes a broad swath of evidence ranging from evidence from a related indication, such as a new indication for an already approved therapy or one adequate and well-controlled trial from two related, but unapproved indications, to strong mechanistic or pharmacodynamic evidence, such as when the underlying biology is well understood and accepted, and the drug is shown to interact with the pathway or biology.

Placebos and the Placebo Effect

A placebo is a substance or treatment that appears real but has no active ingredient or therapeutic benefit. It might be a pill that looks like a real pill but does not contain any drug at all. Sometimes, these are referred to as sugar pills. Or it may be an injection that does not contain any drug. Placebos are used to make a better assessment of a drug's effect and safety, and are often used as a control group in clinical studies.

The placebo effect occurs when a patient responds positively to a placebo, i.e., demonstrates improvement in the absence of an active treatment. The exact mechanisms of the placebo effect are not known, but they are certainly related to a mind-body connection. A nocebo effect is the opposite of the placebo effect, where negative expectations about a treatment lead to adverse events, even when the treatment is a placebo.

Today, the gold standard for an adequate and well-controlled clinical trial is the randomized, placebo-controlled trial, in which patients are randomly assigned to different groups, one of which receives a placebo, and then assigned to their treatment. Typically, no one involved in the study, neither the doctor, staff, patient, nor anyone at the drug company, knows what treatment the patient is receiving; the study is said to be blinded, double-blind in this case. Patients are followed for a period and assessed for an endpoint throughout the trial. At the end of the trial, the data are analyzed and the results reported. There are various modifications to this design, such as blocking, in which patients are also stratified by factors such as sex or disease state. This is intended to provide additional control over treatment assignment and minimize bias.



Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

A key component in assessing efficacy is the endpoints in a clinical study. Endpoints are predefined, specific items that will be measured in a study, and are defined by their importance. Primary endpoint(s) are the main outcomes used to assess a trial. Failure on the primary endpoint often means a failed trial. A good primary endpoint will be clinically meaningful, measurable, clearly defined, and unbiased. Examples of primary endpoints are overall survival in cancer drug trials or time to first heart attack in blood pressure medication trials. The number of patients enrolled in a study, called the sample size, is often based on the variability of the primary endpoint in the population of interest in the absence of the drug. Highly variable endpoints require larger patient numbers in the trial.

Endpoints in clinical trials are clinically meaningful. A recent development is the use of Patient Reported Outcomes (PROs) as endpoints in clinical trials. PROs assess quality-of-life measures for patients and can be used to support label claims. Their results are based directly on patient data, without any physician or other party assessment. Examples of PROs include pain assessment, fatigue, anxiety, depression, and physical assessment of daily activity and mobility. PROs may be collected using wearables, other electronic measures, or by self-reported patient diaries. PRO endpoints may be used as primary or secondary endpoints in clinical trials. Primary PRO endpoints used as the basis for drug approval include:

- Aimovig (erenumab) used patient-reported daily migraine occurrence and severity for migraine prevention;
- Veozah (fezolinetant) used patient-reported daily hot flash occurrence and severity for hot flash prevention in women with menopause;
- Linzess (linaclotide) used patient-reported improvements in irritable bowel symptoms and number of daily bowel movements for prevention of IBS with constipation; and,
- Since there is no measure of "dryness" in the eye, Restasis (cyclosporine ophthalmic emulsion) used patient-reported ocular discomfort and dryness symptoms as the basis for approval.

Important Endpoints in Oncology

Progression-Free Survival: how long a patient lives without the tumor growing or spreading to other parts of the body (metastasis).

Overall Survival: how long a patient lives regardless of cause of death; OS is the gold standard for oncology trials.

Objective Response Rate: the percent of patients who respond to the drug.

Duration of Response: length of time a tumor responds to the drug without it growing or spreading to other parts of the body.

Complete Response: disappearance of all tumors.

Partial Response: >30% decrease in tumor size and no new lesion.

Stable disease: no change in tumor size.

Progressive disease: >20% increase in tumor size from nadir or new lesions observed.

Patient-Reported Outcomes are increasingly being used as secondary endpoints in clinical trials. Common PROs in oncology studies are assessing symptom burden like pain and fatigue, mobility, emotional burden like anxiety and depression, and overall quality of life.

These concepts are illustrated in Figure 67.

Today, PROs are the foundation of patient-centered drug development, which holds that approval should be based on metrics that matter to patients. For example, a cancer drug may improve survival by a few months, but if the drug reduces the patient's mobility or quality of life, then patient advocates argue that the drug is not useful. PROs, like any other endpoints measured in a clinical trial, must be reliable, validated, and sensitive to change. Further, the PRO must generally be used in the population for which it was intended. For example, a PRO developed for children might not be valid in adults. The FDA has issued guidance on the use of PROs in clinical trials and in support of label claims; the reader is referred to that guidance for more details [224].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

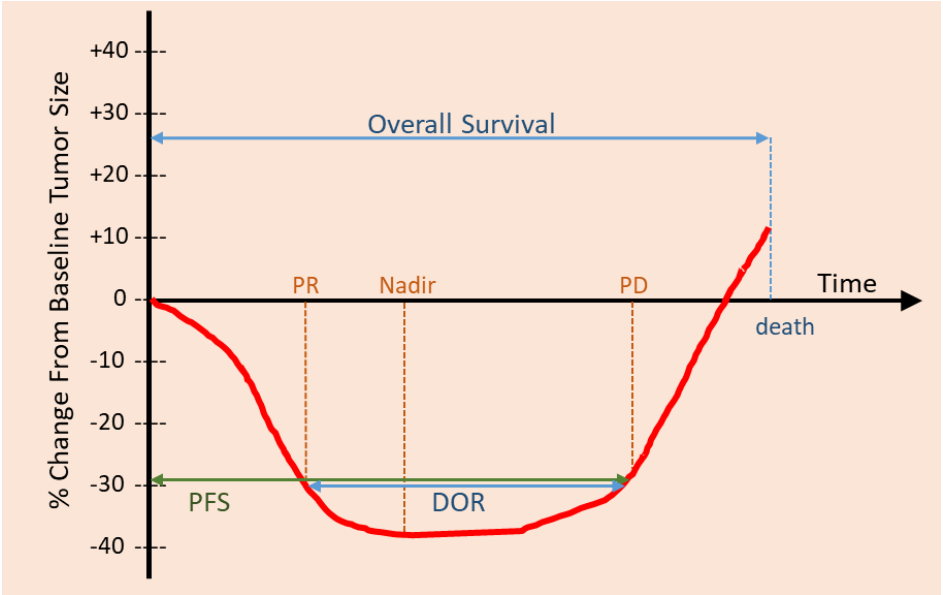


Figure 67: Representative tumor growth profile illustrating the main endpoints in oncology. This patient demonstrates a partial response (PR), having a ~38% decrease in tumor size at the nadir, but after some time, the tumor starts to regrow (progressive disease, PD) and continues growing until the patient dies.

Sometimes, multiple endpoints are combined to form a composite endpoint. Composite endpoints may be used when the endpoints occur infrequently. A classic example of a composite endpoint is MACE, which stands for Major Adverse Cardiac Events, and is the occurrence of a heart attack, stroke, or cardiovascular death in the group of interest. Each of these events may occur infrequently, but together they can raise the event rate to yield a meaningful estimate. Nevertheless, clinical trials using MACE as an endpoint still typically require thousands of patients.

Some trials use multiple endpoints as their primary endpoint. For example, fezolinetant, a drug used to treat hot flashes in menopausal women, in the Skylight-1 pivotal trial (NCT04003155), had four coprimary endpoints: mean change in frequency of moderate-to-severe vasomotor symptoms, i.e., hot flashes, from baseline to weeks 4 and 12, and mean change in severity of moderate-to-severe hot flashes from baseline to weeks 4 and 12 [225]. In addition to the primary endpoints, most studies have secondary and exploratory endpoints. Secondary endpoints are ancillary endpoints used further to assess the safety and effectiveness of a drug. For example, the fezolinetant Skylight-1 trial had multiple secondary endpoints, the primary one being patient-reported assessments of sleep disturbance after 12 weeks of therapy. Menopausal women often report difficulty sleeping because hot flashes occur during the night. Knowing whether their sleep improves is further evidence of fezolinetant reducing the frequency and severity of their hot flashes. Exploratory endpoints are not the outcomes of interest and may be hypothesis-generating or provide supporting information. Example exploratory endpoints may include assessing a drug's pharmacokinetics in the population, examining changes in biomarkers, or further evaluation of patient-reported outcomes. In Skylight-1, exploratory endpoints were measurement of fezolinetant and its primary metabolite concentrations in plasma, biomarker assessment, and additional patient-reported outcomes on quality of life.

All the details of a clinical trial, such as the patient population, treatments, endpoints, statistical analysis, and reporting guidelines, are documented in a protocol. Protocols for pivotal studies are often sent to the FDA for review and approval. The FDA requires companies to post their clinical trials, including results, on clinicaltrials.gov; however, these are essentially Cliff's Notes or

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

abbreviated versions of the studies. Some companies publish entire protocols in journals such as BMJ Open to promote transparency. Because of the variability in protocol formats across companies, harmonized guidelines, like SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), have been published to make protocols more uniform [226].

More on the design and analysis of clinical trials can be found in Appendix 2.

What is Confirmatory Evidence?

Historically, approval of a new drug was based on results from randomized clinical trials. Since 1998, FDAMA has allowed approval based on one adequate and well-controlled trial, along with confirmatory evidence. Since 2023, the FDA has provided guidance on what types of data provide confirmatory evidence [223]. When reviewing applications for approval, the FDA considers both the quality and quantity of evidence. Further, sponsors are encouraged to discuss with the FDA what confirmatory evidence will be provided, and that will be sufficient for approval.

The FDA guidance on confirmatory evidence provides the following examples of confirmatory evidence:

1. Clinical evidence from a related indication. For example, a drug submitted to treat psoriatic arthritis has been shown to work in rheumatoid arthritis.
2. Mechanistic or pharmacodynamic evidence. For example, the underlying biology of the disease is well understood, and the drug's mechanism of action has been shown to target the disease pathway.
3. Evidence from a relevant animal model. For example, when the drug is an antibiotic and there is a proven relevant animal model of the disease in which the drug is effective.
4. Evidence from other members of the same pharmacological class. For example, blockade of β-adrenergic receptors has been shown to decrease blood pressure, and the drug is proven to be an β-blocker.
5. Pharmacokinetic-Pharmacodynamic modeling and simulation. For example, Aristada (aripiprazole lauroxil) used modeling and simulation to support the approval of a new dose strength and new dosing regimen without the need for additional clinical trials [227].
6. Use of natural history or real-world evidence. For some rare diseases, running a control group in a confirmatory trial is unethical, or there are too few patients with the disease of interest to enroll in a clinical trial of sufficient size to make rigorous conclusions afterward. In this case, natural history data might be compared against the active treatment arm in a clinical trial. For example, Exondys 51 (eteplirsen) was approved in 2016 for the treatment of Duchenne muscular dystrophy. In a controversial approval, Exondys 51 failed to meet its primary endpoint after 1 year of treatment but reached statistical significance at 3 years after treatment initiation using an external control arm of 13 patients (Figure 68). The issue was not the external control but the failure to meet its endpoint, which was specified as 1 year in the protocol. The FDA also used the increase in dystrophin (a protein that strengthens and protects muscle fibers) in skeletal muscle as a surrogate marker of benefit to patients.

The Animal Rule

For some drugs, it is unethical or infeasible to conduct clinical trials in humans. For example, with a drug to treat nerve gas toxicity or a drug to treat anthrax exposure, it would be unethical to administer nerve gas or anthrax to a volunteer and then test the effects of the drug on them. Under these conditions, the FDA can approve a drug based on results from animal studies using the so-called Animal Rule [228]. In these circumstances, "*FDA may grant marketing approval based on adequate and well-controlled animal efficacy studies when the results of those studies establish that the drug is reasonably likely to produce clinical benefit in humans*."

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

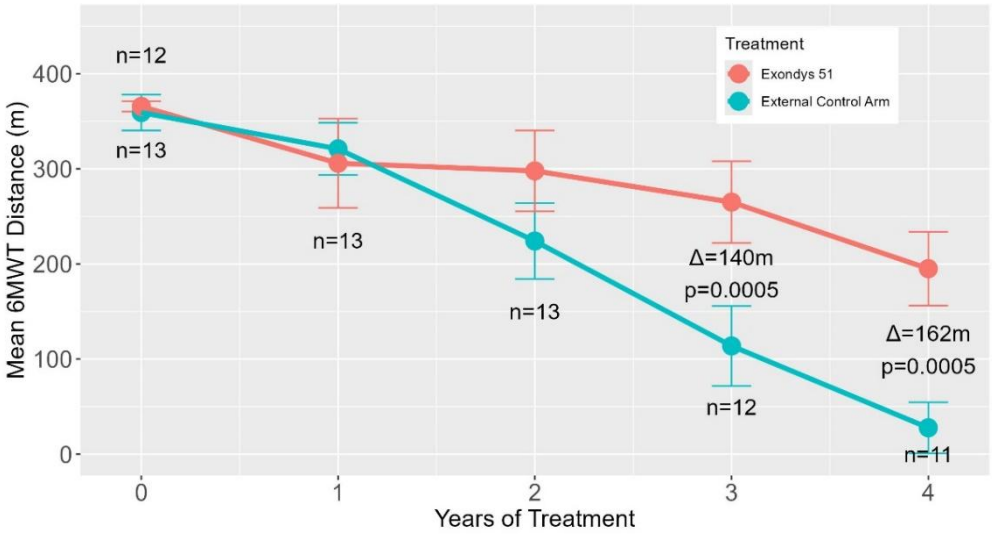


Figure 68: Mean distance walked in 6-minutes (6MWT) in Duchenne's muscular dystrophy patients treated with Exondys 51 for 4 years compared to an external control arm of patients not enrolled in the system.

The Animal Rule has four conditions:

1. The pathophysiological toxicity or mechanism of the illness is well understood.
2. The effect is demonstrated in one or more species and is predictive of the response in humans.
3. The animal study endpoint is clearly defined and related to clinical benefits in humans, like survival.
4. The PK and PD of the product are such that animal exposure can be used to choose an effective dose in humans.

Under these conditions, there is a reasonable expectation that animal results can be extrapolated to humans.

Examples of drugs approved under the animal rule include:

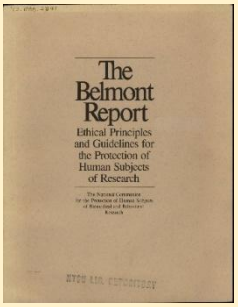
- Pyridostigmine for the prophylaxis against soman nerve gas poisoning (approved 2003);
- Cyanokit (hydroxocobalamin) for the treatment of cyanide poisoning (approved 2006);
- Equine Botulism Antitoxin Heptavalent (A, B, C, D, E, F, G) for the treatment of symptomatic botulism (approved 2013);
- And numerous agents against anthrax, plague, and smallpox disease.

One example of the Animal Rule is Cyanokit's approval. For this approval, a dog model was used to establish efficacy, along with uncontrolled n-of-1 case studies of patients treated with Cyanokit after smoke inhalation or accidental or intentional cyanide ingestion was used as evidence of effectiveness. These were deemed sufficient evidence of effectiveness that approval was granted in 2006.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Is It Ethical to Use Children in Clinical Trials?

Studies done with human volunteers must adhere to certain ethical standards. This seems obvious today, but it wasn't always the case. Before the Wiley Act of 1906, there were no ethical standards. After World War II, when Nazi "scientists" were placed on trial for war crimes and crimes against humanity, the first ethical guidelines were issued. In their final ruling, the judges issued 10 research guidelines for the future, designed to protect research subjects. Although they did not carry the weight of law, these guidelines, known as the Nuremberg Code, were the first to introduce the concept of informed consent – that is, research participants would be informed of the risks and benefits of the research, and it would be their choice whether or not to participate in the study. Participation was voluntary. Other guidelines stipulated that the research would be conducted by qualified personnel, that the study would be based on animal testing and knowledge of the disease being studied, and that the experiment would be halted if there was a risk of injury to the participants.



After the thalidomide disaster, the World Medical Association issued the Declaration of Helsinki in 1964. While the Nuremberg Code was meant to protect research subjects, the declaration focuses on the ethical responsibilities of physicians. Many of the concepts outlined in the Nuremberg Code are reflected in the declaration, including the principles of informed consent and the right of volunteers to make their own decisions regarding participation in the study. There are some clarifications as well. The declaration states that the rights of the volunteer outweigh any benefits to society and that, in certain circumstances, a placebo group or no-treatment arm may be used. Furthermore, whereas the Nuremberg Code requires informed consent in all situations, the Declaration recognizes that informed consent is not possible in some situations and that a proxy may be used to provide consent. The declaration is morally binding for physicians and is intended to supersede local or national laws. Since its inception, the declaration has been revised and updated numerous times, most recently in 2024. Because of some of these changes, the declaration is not universally accepted. Due to the introduction of new language regarding the use of placebos, the 2000 revision was rejected by the FDA, which subsequently discontinued the declaration altogether in 2006, replacing it with Good Clinical Practices when those guidelines were issued in 2008. Similarly, the European Union only uses the 1996 version.

There are also other ethical guidelines, and researchers sometimes disagree over which to follow. In response to the Tuskegee Syphilis Study, wherein hundreds of African-Americans who had syphilis were monitored for decades, never telling the participants they had the disease and simultaneously being denied treatment when treatment became available, the United States established in 1974 the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The commission issued the Belmont Report, which established ethical principles for human research. It contained three guiding principles: autonomy and respect for persons, beneficence (not harm), and justice (benefit-to-risk analysis). This report served as the basis for the Common Rule, issued by the Department of Health and Human Services in the United States, which is now followed by 20 US government agencies, including the FDA. Furthermore, the International Conference on Harmonisation issued Good Clinical Practice (GCP) guidelines in 2008. To further complicate matters, some local rules and regulations may be in place to protect human subjects in clinical studies.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Is It Ethical to Use Children in Clinical Trials?

The protocol often outlines the ethical guidelines that a particular study adheres to. For example, one such statement reads: "*This study will be conducted in accordance with the protocol and with the following: Consensus ethical principles derived from international guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences international ethical guidelines, applicable ICH GCP guidelines, and applicable laws and regulations*." Further publication in journals such as the New England Journal of Medicine or JAMA requires an ethics statement confirming that the trial complied with relevant ethical guidelines.

There is a need for approved drugs for children because, before 1997, there was no mandate for sponsors to study a drug in children to obtain approval. Today, most drugs are developed for adults and children. Returning to the central question, is the use of children in clinical trials an ethical practice? This controversial topic was raised in the context of the pivotal trials used to assess the mRNA COVID vaccines developed in the early 2020s, which employed randomized, placebo-controlled, double-blind study designs (one such study is described here).

There are many issues and no straightforward answer, but two relate to benefit and informed consent. Under the Common Rule, unless the risks of the treatment are minimal or no more than a minor increase, the drug must offer a likely therapeutic benefit with a favorable benefit-to-risk ratio to enroll children in a clinical trial. In the case of the COVID-19 vaccine, this condition appears to have been met based on clinical trials in adults. The second issue is that children cannot give informed consent; however, many ethical guidelines allow their parents to give informed consent on their behalf. Some have argued that parents cannot give informed consent when their child is involved because of their inherent conflict of interest. If you had a child with a rare disease that might kill them at a very young age, if a company came to you and said they had a potential drug that might treat their condition and prolong their life, if you enrolled in their clinical trial, would you turn it down? Regardless of the risks? This is where the uncertainty lies. Depending on whether you think a parent can objectively weigh the risk-benefit of an experimental drug, it leads to whether you think it's appropriate to study an experimental drug in children.

And what about the use of placebo treatments in clinical trials involving children? One review paper from 2021 reported that of the 266 products under pediatric development between 2012 and 2018, 25% used a placebo control in children, although half of these involved using a placebo given in combination with an already approved therapy [229]. Pediatric patients assigned to a placebo arm will not benefit from the medication; therefore, comparing a new drug to a placebo is unethical if an already approved drug exists for the illness. The new drug must be compared to the approved drug. Furthermore, the use of placebos in children must be minimized. Perhaps, instead of randomizing 1:1 between drug and placebo, a 2:1 ratio could be used. Suppose a futility analysis indicates that the drug is superior to a placebo early on. In that case, it is unethical to continue the study as is, and placebo patients could then be reassigned to the experimental drug and followed up on for safety and efficacy.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Real-World Data and Evidence. This is another one of those topics that cannot be covered in its entirety and a high-level overview will be presented instead.

The 21st Century Cures Act of 2016 was designed to increase the efficiency of drug development, bringing products to the market faster. One of the provisions in the act was a framework for using real-world evidence (RWE) to support drug approval under section 505(c) of the FFDCA [230]. The FDA distinguishes between data and evidence, which are not the same. Real-world data (RWD) is a broad term that the FFDCA defines as *"data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources, such as electronic health records (EHRs), claims and billing activities, product and disease registries, patient-generated data from in-home settings, and data from other sources, such as mobile devices"* [231]. RWE uses real-world data to provide clinical evidence about the potential benefits and risks of a drug. Under the FFDCA, the FDA must evaluate whether RWD has the potential to generate RWE to help support approval of a drug for a new indication.

The first FDA approval of a drug for a new indication based solely on RWD was granted to Astellas for Prograf (tacrolimus). Tacrolimus is an older drug used in solid-organ transplants to prevent rejection, specifically in the liver, kidneys, and heart. Astellas submitted electronic health record data from all lung transplants in the US. They showed that outcomes in patients treated with Prograf were significantly better than in those who did not receive it. *No clinical studies were done to support this application.* The FDA concurred and approved Prograf for use in lung transplant patients in 2021.

Regardless of whether RWD can be used to generate RWE, RWD has its uses in improving clinical trial efficiency. RWD can be used to:

- help generate hypotheses for future clinical trials,
- help identify biomarkers that might be useful and are associated with a population,
- help improve the selection of inclusion/exclusion criteria for clinical trials,
- help identify the statistical distribution of a biomarker, like its mean and standard deviation, in a population, and
- help identify patient factors that might influence the outcome of a disease.

One example of using RWD to improve the efficiency of a drug development program is the use of a biomarker to determine whether a patient may respond to treatment with a particular drug. For example, Zelboraf (vemurafenib) is a kinase inhibitor used to treat melanoma in patients with a specific DNA mutation: the BRAF B600E mutation. Patients must be screened using a diagnostic test to ensure they have the mutation before they can be prescribed the medication. Zelboraf is specifically not recommended for patients without the mutation. About 50% of patients with melanoma have a BRAF mutation, and of those, about 90% have the V600E mutation. The V600E mutation was identified using RWD, and vemurafenib was specifically designed to target it.

Another use of RWD is as an external or historical control arm in a clinical trial. In a controlled clinical trial, the comparator arm is often a placebo treatment group. In rare diseases, assigning some of your patients to a placebo arm is wasteful. Those patients are valuable because they are often hard to find and may be few in number. Using natural history data, in other words, how the disease progresses over time without treatment, may be used as a comparator arm in a clinical study [232]. Patients receiving the treatment who are enrolled in the clinical trial are compared to the external control group, which consists of individuals not enrolled in the trial and not receiving the treatment of interest, whose data were collected at the same or different relative point in time, or possibly in another setting. The results from the external control group must be extrapolated to the clinical trial setting. A major issue with an external control group is its representativeness relative to the patients enrolled in the clinical trial. One advantage to using an external control group is that all the patients can be enrolled in the treatment of interest arm in the clinical trial, improving the ability to estimate its overall effect and safety. The disadvantage is that external control groups tend to overestimate the therapeutic benefit of the treatment of interest.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 12

Efficacy Results from Tivozanib Phase 3 Confirmatory Trial Against Sorafenib

Endpoint	Tivozanib (n=175)	Sorafenib (n=175)
Progression Free Survival (PFS)		
Events, n (%)	123 (70)	123 (70)
Progressive Disease	103 (59)	109 (62)
Death	20 (11)	14 (8)
Median (95% CI), months	5.6 (4.8, 7.3)	3.9 (3.7, 5.6)
Hazard Ratio* (95% CI)	0.73 (0.56, 0.95)	
p-value	0.016	
Overall Survival		
Deaths, n (%)	125 (71)	126 (72)
Median (95% CI), months	16.4 (13.4, 21.9)	19.2 (14.9, 24.2)
Hazard Ratio* (95% CI)	0.97 (0.75, 1.24)	
Objective Response Rate % (95% CI)	18 (12, 24)	8 (4, 13)
Median duration of response in months (95% CI)	NE (9.8, NE)	5.7 (5.6, NE)

CI, confidence interval; NE, not evaluable. Table reprinted from the Fotivda label [233].

* A **Hazard Ratio** is the instantaneous risk of the event occurring in the group of interest relative to the comparator group. A hazard ratio < 1 indicates improvement, whereas > 1 indicates worsening. If the confidence interval (CI) contains 1, there is no difference between groups.

An example of a drug approved using an external control arm is Koselugo (selumetinib), a kinase inhibitor used to treat pediatric patients 2 years or older with neurofibromatosis Type I (NF1) who have inoperable neurofibromas. These noncancerous tumors grow along nerves throughout the body. NF1 is a genetic disorder that typically begins in childhood and affects 1 in 3,000 to 4,000 people worldwide. Although not lethal, NF1 can be painful, and patients experience chronic pain throughout life. Because it is a rare disease, it was difficult to find and recruit patients for clinical trials testing selumetinib. Approval of Koselugo was based on the results of the SPRINT Phase II Stratum 1 study, a Phase 2 trial in 50 NF1 patients aged 3.5 to 17 years. AstraZeneca used the results from two external control groups. One was a natural history study in 92 patients, and the other was the placebo arm of a Phase 2 study in NF1 patients using a different drug, tipifarnib, which had 29 patients (NCT00021541). Although no formal statistical comparison was made between the trial group and the external control groups, the FDA concluded that the probability of spontaneous regression of neurofibromas was small and that the regression seen in the selumetinib patients was likely drug-related. The drug was approved in 2020.

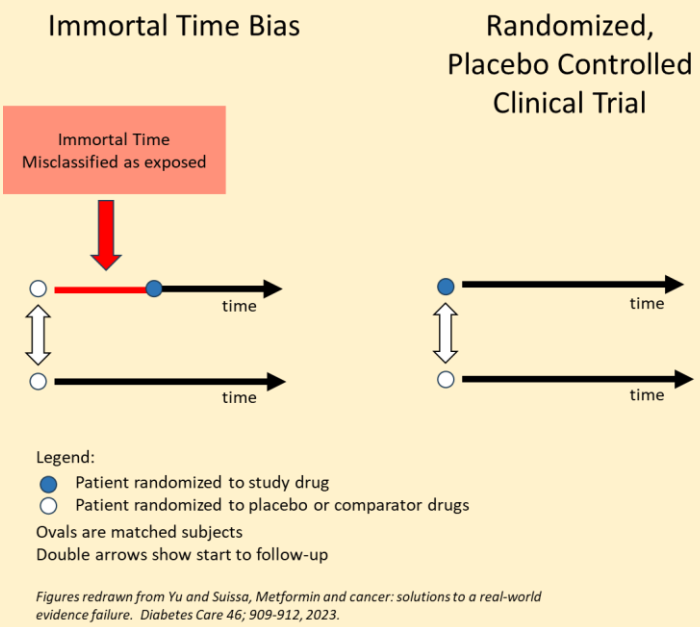
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Does Metformin Prevent or Cure Cancer? An Example of Unintentional Bias?

Metformin is widely used to treat Type 2 diabetes. In studies to determine metformin's mechanism of action, it was shown to inhibit AMP-activated protein kinase (AMPK), an enzyme that regulates downstream pathways critical for cell growth and proliferation. Preclinical studies have demonstrated that metformin inhibits tumor growth, particularly in breast cancer cells. On this basis, an epidemiological study was done that showed patients with Type 2 diabetes being treated with metformin were less likely to develop cancer than those not being treated with metformin [234]. Other additional meta-analyses suggested that metformin decreased cancer risk by 30% to 50% [235].

Prospective clinical trials have been less convincing, however. In the most comprehensive study to date, a randomized, double-blind, placebo-controlled trial in 3649 Type 2 diabetics with breast cancer who were to be followed for 5 years, metformin as an adjuvant to cancer therapy did not affect survival, and the study was stopped early for futility (basically, the study was stopped because metformin had no chance of improving survival) [236]. Why, then, do placebo-controlled, randomized clinical trials (RCTs) not show an effect, but observational studies do? The original epidemiological studies were retrospective, observational studies, which can be biased and lead to erroneous conclusions.

One bias that can occur, and has been suggested as being the reason for the original findings, is something called immortal-time bias (ITB) [237, 238], which occurs when an individual is exposed to the treatment after the start of the observational period, such that there is a period of time before treatment in which they were not exposed to the treatment (see figure). This can result in an overestimation of the event rate in the control group, an underestimation in the treatment group, or both. It has been argued, however, that the magnitude of the bias introduced by ITB cannot account for the effect observed in meta-analyses with metformin, i.e., ITB does not account for a 30% to 50% increase in survival.



Figures redrawn from Yu and Suissa, Metformin and cancer: solutions to a real-world evidence failure. Diabetes Care 46; 909-912, 2023.

So why have the clinical trials failed? Clinical trials often fail for several reasons. One is that the wrong population was studied. Maybe breast cancer is not the right cancer type for metformin to have an effect. Maybe the dose was too low? Clinical trials with metformin have used the clinical dose used to treat diabetes – maybe higher doses are needed for an antitumor effect [239]. And maybe metformin does not affect tumor growth, but may still be useful as a cancer-preventative agent. Several cancer prevention trials are ongoing. For now, the jury is out regarding metformin's role in preventing and/or curing cancer.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Blincyto (blinatumomab) is an example of a study that used an external control group to compare effectiveness. Blinatumomab is a bispecific antibody for use in acute lymphoblastic leukemia (ALL). The clinical trial supporting approval was an open-label, single-arm study with no control group. The sponsor, Amgen, used an external control arm of 694 ALL patients from the US and the E.U. The primary endpoint of the study showed that 42% of patients treated with blinatumomab achieved complete leukemia regression, compared with 24% in the control arm. The FDA agreed that this provided evidence of efficacy and granted accelerated approval in 2014. The FDA required a post-marketing study to confirm efficacy, using an internal control group.

It is to be expected that RWD and RWE will be used more widely. The volume of RWD continues to grow, and regulatory agencies are expected to be more receptive to submissions that include such data. Furthermore, machine learning can be used to merge, preprocess, and analyze diverse real-world datasets, thereby helping identify patterns within the data. This will be a huge boon, as it is currently quite challenging to merge and curate datasets from different sources.

The Clinical Studies Section of the Package Insert

Efficacy results from clinical trials are reported in Section 14 Clinical Studies of the package insert. As an example of how efficacy is reported in this section of the label, let's examine the label for Fotivda (tivozanib), a cancer drug approved in 2024 for the treatment of adult patients with relapsed or refractory renal cell carcinoma (a type of cancer of the renal tubules) after failing two or more prior therapies. A cancer drug was chosen because the endpoints are easy to understand. In a confirmatory trial, tivozanib went head-to-head against sorafenib, another already approved drug for the same indication. This confirmatory trial, called TIVO-3, was a randomized (1:1), open-label study (patients, doctors, and the sponsor knew what therapy each patient was randomized to). Patients were followed for 6-months and assessed for progression-free survival (PFS) as the primary endpoint. Secondary endpoints were overall survival (OS), objective response rate (ORR), and median duration of response. The results from the trial are presented in Table 12. More patients responded to tivozanib overall. Tivozanib improved median progression-free survival by about 6 weeks and median overall survival by almost 3 months. Shown on the table are p-values and corresponding confidence intervals. The regulatory review for tivozanib stated the basis for approval was the TIVO-3 clinical trial. The reviewer concluded that tivozanib showed a significant improvement in PFS compared with sorafenib. Although the OS was similar across groups, the ORR was higher in the tivozanib group, suggesting tivozanib was more effective than sorafenib.

Safety and Adverse Events

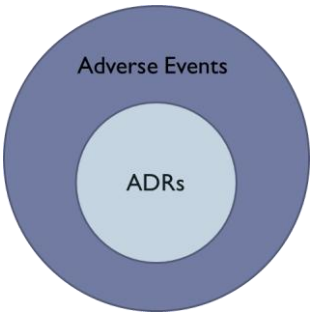
Safety has been a requirement for marketing approval by the FDA since the original FFDCA in 1938, and if you like acronyms, then you will love this section on adverse events, or AEs, as they are often called.

AE. ADR. SUSAR. Say What?

We have all experienced an adverse event after taking a drug. Perhaps you have experienced diarrhea or an upset stomach after taking antibiotics for an extended period. Maybe you got drowsy after taking some cold medicine. Or maybe you had something potentially more dangerous, like experiencing lightheadedness after taking your blood pressure medicine. The statistics of AEs are stunning. Researchers at Utrecht University in the Netherlands estimated that 3.5% of all hospital admissions in Europe were due to an adverse drug reaction (ADRs), which is an unwanted, undesirable effect that is a direct result of taking a drug in routine clinical use [240]. It has been further estimated that 5% of all hospitalized patients in the EU will experience an ADR at some point during their hospitalization and that ADRs account for approximately 200,000 deaths each year in the EU. The statistics are no better in the US, where 6% of hospital admissions were due to ADRs and 3% of ADRs occurred while hospitalized [241]. The cost of managing an ADR can also be quite high. One estimate puts the cost to the US health care system at over $30 billion per year [242]. Another study estimated that a single ADR in a hospital can add approximately (adjusted for inflation in 2024) $US 26,000 to the hospitalization cost in a non-intensive care unit ward and about $US

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

39,000 in the intensive care unit [242].



An AE event is formally defined as any untoward event associated with the use of a drug in humans, *whether or not considered related*; causality is not a factor. ADRs are a subset of adverse events, defined as any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product that has a causal relationship with this treatment. In other words, an ADR is a direct result of the drug, while AEs may or may not be. For instance, suppose you slip on a banana peel while walking down the sidewalk hours after you took a new experimental medication for a runny nose because you were looking at your phone while walking. That is an AE. Suppose you slipped on a banana peel while drowsy after taking the medicine and did not see it. That would be an ADR. If you hurt yourself enough to require a trip to the hospital, that would be a serious adverse event (SAE), which is defined as any untoward medical event that at any dose:

- results in death,
- is life-threatening,
- requires inpatient hospitalization or causes prolongation of existing hospitalization,
- results in persistent or significant disability/incapacity,
- may have caused a congenital anomaly/birth defect, or
- requires intervention to prevent permanent impairment or damage.

If an SAE can be directly attributable to the drug, then that would be a Serious Adverse Reaction (SAR), but if the SAE was unexpected, like if it was never seen before and this was the first case report, then this would be considered a SUSAR, a Suspected Unexpected Serious Adverse Reaction, which if you like acronyms are usually B-A-D, i.e., bad.

Drug-Related Adverse Events in the News

SUNDAY EXPRESS

WIN UP TO £2,000 IN OUR BUMPER PUZZLE SPECIAL

PLUS FREE ICE CREAM TODAY AT LONDIS

Exclusive: Fresh calls for review into safety of heart drug as 20,000 suffer side effects

STATINS LINKED TO 227 DEATHS

Crown Jewel for baby Charlotte

Terror chiefs fear UK attack could come any day

TORONTO STAR

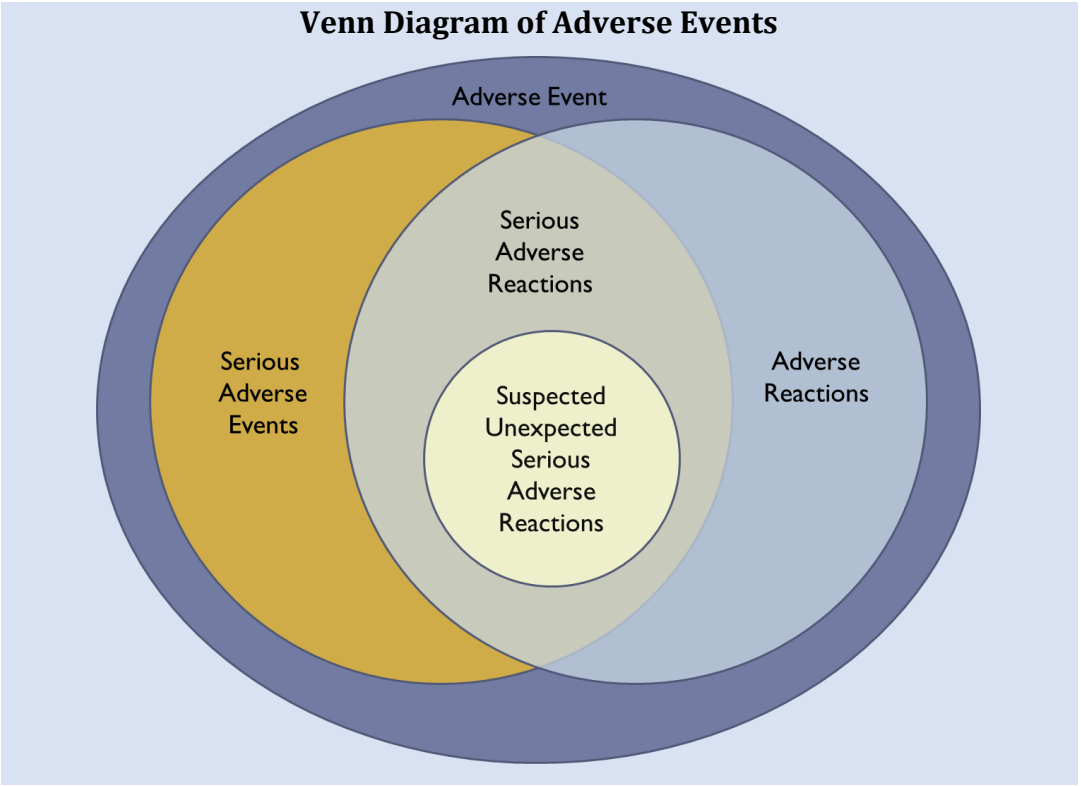
A wonder drug's dark side

Hundreds of thousands of teen girls in Canada have safely taken Gardasil, a vaccine shown to prevent HPV. But a Star investigation has found that since 2008, at least 60 Canadians experienced debilitating illnesses after inoculation. Patients and parents say the incidents point to the importance of full disclosure of risks

Types of Adverse Events

There were two basic types in the original classification of adverse events: Type A, Type I, or expected AEs, and Type B, Type II, or unexpected AEs. For this book, we will use Type A and B to avoid confusing Type I or Type II errors. Type A AEs are often called side-effects. They are expected based on the type of medication. Benzodiazepines cause retrograde amnesia when taken; anticholinergics can cause motion sickness or dry mouth, and warfarin may cause bleeding issues. These are all examples of Type A AEs. Most Type A AEs result from the drug acting at off-target receptors, which

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



are not at the desired location (e.g., acting in the brain when the primary target is the heart) or by a greater than desired magnitude of effect at the site of action. On the other hand, rhabdomyolysis (a serious muscle injury) from statin use, anaphylaxis to penicillin, or a rash from sulfa antibiotics are not expected. These are Type B AEs. Type A AEs are the most common and are related to the drug's mechanism of action. They are predictable and have low mortality. Type B events are uncommon, not related to the mechanism of action of the drug, and may be life-threatening. For a Type A event, you might lower the dose or withhold the drug and then restart at a later time. For a Type B AE, you would stop therapy and probably never restart it. When a doctor's office asks you if you have any drug allergies, this question usually refers to whether you have ever had a Type B reaction to a drug.

Later, additional categories were added to the schema. Chronic reactions (Type C) are uncommon AEs that are dose- and time-dependent and related to cumulative dose. An example of this might be doxorubicin-induced heart failure (one study reported that 26% of patients taking doxorubicin might experience congestive heart failure when their cumulative dose exceeds 550 mg/m^2). Type D reactions are uncommon, usually dose-related, and are delayed after initiation of the drug. An example of this might be the tardive dyskinesia patients taking typical dopamine-blocking antipsychotic medication may eventually develop. Type E reactions occur after stopping the drug and are related to drug withdrawal, such as the rebound insomnia that occurs after repeated use of sleeping agents. Type F reactions are unexpected drug failures, such as if a woman gets pregnant while taking an anticontraceptive medicine.

Risk Factors for AEs and ADRs

Not everyone will have an AE or ADR. Many risk factors predispose someone to an event, and attribution to a single factor is rare; predisposition is multifactorial, with some factors being consistent across all drugs and others being drug-specific [243]. Patient-specific factors that might affect the occurrence of ADRs include:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

- Age extremes: the young and old are particularly prone to ADRs.
- Gender: females are more prone to ADRs than males.
- Multiple drugs disproportionately increase the risk of ADRs. For example, patients receiving corticosteroids and non-steroidal anti-inflammatory drugs are 15-times more likely to develop peptic ulcer disease.
- Disease state. Renal or hepatic impairment may increase the risk of an ADR.
- Past history of an ADR. Drug allergies can lead to cross-sensitization to other drugs.
- Obesity increases the risk of an ADR.
- Genetic factors. Certain polymorphisms may increase the risk of an ADR.

There are also drug-related factors that might affect the occurrence of an ADR:

- Drug dose and frequency: higher total doses increase the likelihood of an ADR.
- Route of administration: Intravenous drug administration is associated with a greater risk of ADRs than the oral route of administration.
- Change in formulation or drug product: A change in formulation could result in a SUSAR not related to the drug, or, for biologics, a change in process could change the product's glycosylation pattern, resulting in a change in antigenicity.

Diethylstilbestrol was used in the 1950s to prevent miscarriages but was found decades later to cause vaginal cancer in the daughters of pregnant mothers who took the drug.

Image courtesy of Diethylstilbestrol DES, Journal of a DES Daughter (https://diethylstilbestrol.co.uk).

Assessing Causality

Assigning causality to a drug can be difficult. Consider these two examples:

1. An obese 65-year-old male smoker who is hypertensive and has poorly controlled type 2 diabetes is treated with a cholesterol-lowering drug and has a heart attack 8 days after starting therapy. Was this drug-related?
2. A pregnant woman takes a drug to prevent miscarriage, and 14 years later, her daughter develops vaginal cancer. Was this drug-related?

In the first example, it seems likely to be drug-related, but in the second example, it does not seem related. But consider this. In the first example, the man's parents both lived into their 90s and had no history of heart disease. He only took a single dose of the drug because it gave him a headache, and on the day of the attack, he was outside working in his garden for 5 hours in the hot sun. Now, it seems less likely. The second example is drug-related. Diethylstilbestrol (DES) was an estrogen prescribed to women in the 1950s to prevent miscarriages. A total of 5 to 10 million women in the US were prescribed DES because doctors thought it was safe. In 1953, its use was discontinued because it was shown to be ineffective. In 1971, a retrospective epidemiological study linked DES to a rare form of vaginal cancer in the daughters of DES mothers. It is clear from these examples that causality determination can be tricky.

Pharmacovigilance (PV) is the department in the pharmaceutical industry responsible for collecting, detecting, assessing, monitoring, and preventing adverse effects associated with pharmaceutical products. Some of the basic criteria used to assess causality are:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 13 WHO-UMC Causality Assessment System	
Causality Term	**Assessment Criteria**
Certain	• Clinical event or laboratory test abnormality that occurs at a plausible time in relation to drug administration • Cannot be explained by underlying concurrent disease, other drugs or chemicals • Response to the withdrawal of the drug is clinically plausible • The event is definitive pharmacologically or phenomenologically • If necessary, a rechallenge is necessary
Probable/ Likely	• Event or laboratory test abnormality with a reasonable time relationship to drug intake • Unlikely to be attributed to disease or drugs • Response to withdrawal is reasonable • Rechallenge not required
Possible	• Event or laboratory test abnormality with a reasonable time relationship to drug intake • Could also be explained by disease or other drugs • Information on drug withdrawal may be lacking or unclear
Unlikely	• Event or laboratory test abnormality with a time to drug intake that makes a relationship improbable (but not impossible) • Disease or other drugs provide plausible explanations
Conditional/ Unclassified	• Event or laboratory test abnormality • More data is needed on the proper assessment needed, or • Additional data under examination
Unassessable/ Unclassifiable	• Report suggesting an adverse reaction • Cannot be judged because the information is insufficient or contradictory • Data cannot be supplemented or verified

- Is the event consistent with the pharmacology of the drug or previous knowledge of the drug's ADRs?
- What is the association between drug administration and the time of the event?
- How medically or biologically plausible is the event?
- Can other causes be excluded?
- Is the patient taking multiple medications? Is the ADR the result of a drug interaction?
- Is it dose-related? Does cessation of therapy cause the ADR to resolve or lessen in intensity? Does the reaction intensity decrease when the dose is reduced?

ADRs are often assigned a causality category. For instance, the World Health Organization, in partnership with the Uppsala Medical Center (WHO-UMC, see Table 13), has issued an assessment scale with six categories: certain, likely, possible, unlikely, unclear, or unclassifiable. Most ADRs are not classified as certain or assessable because most will fall into the in-between categories. Other scales, like the Naranjo ADR Probability Scale, have 10 questions based on the basic causality criteria previously discussed, answered as yes (+1), no (-1), or don't know (0). After adding the individual scores, a composite score is assigned a probability category: highly probable (9), probable (5-8), possible (1-4), or doubtful (0). There are other scales and algorithms, but none are definitive in their assessments.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Severity of ADR

The severity of all AEs or ADRs in any clinical study must be assessed. Severity is frequently a judgment call by the physician overseeing the patient. It is recorded by the physician in the Case Report Form, which is a document used to collect data from participants in a study. In addition to the start and stop date of the AE, which must be recorded, severity is graded as follows:

- Mild: an easily tolerated event that causes minimal discomfort and does not interfere with daily activities;
- Moderate: an event that is sufficiently discomforting and interferes with daily activities; or
- Severe: an event that prevents normal everyday activities.

Oncology studies and many non-oncology studies are also graded using severity scales set forth by the National Cancer Institute. Under this scale, called the Common Toxicity Criteria for Adverse Events (NCI-CTCAE), severity is defined as:

- Grade 1 Mild: asymptomatic or mild symptoms; clinical or diagnostic observations only; no intervention indicated;
- Grade 2 Moderate: minimal, local, or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living;
- Grade 3 Severe or medically significant but not immediately life-threatening: hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living;
- Grade 4 Life-threatening consequences: urgent intervention indicated; or
- Grade 5: Death related to AE.

Other alternatives are the World Health Organization Toxicity Criteria (WHOTC), which is also a 5-point Likert scale. The scales are similar but not the same. For example, the following table compares the CTCAE to the WHOTC for nausea and vomiting:

Toxicity Criteria Scales for Nausea and Vomiting

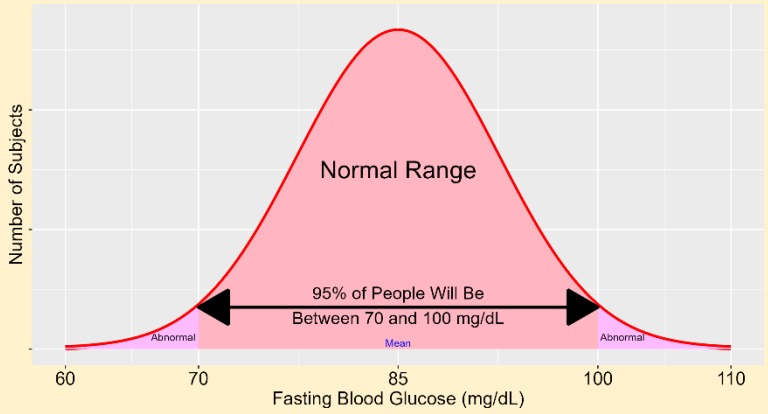
Grade	CTCAE – nausea	CTCAE – vomiting	WHOTC
0	None	None	None
1	Loss of appetite without alteration in eating habits	Intervention not indicated	Nausea
2	Oral intake decreased without significant weight loss, dehydration or malnutrition	Outpatient IV hydration; medical intervention indicated	Transient vomiting
3	Inadequate oral caloric or fluid intake; tube feeding, total parental nutrition (TPN), or hospitalization indicated	Tube feeding, TPN, or hospitalization is indicated	Vomiting therapy required
4	---	Life-threatening consequences	Intractable vomiting

Notice that the CTC splits nausea and vomiting into two scores, whereas the WHO treats it as one syndrome. Furthermore, if medical intervention is required, this is Grade 2 under CTCAE but Grade 3 under WHOTC. The chosen grading scale must be used consistently throughout a drug development program.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

What is Normal?

If you visit the doctor when you're sick, they often ask for blood and urine samples for laboratory tests. The concentrations of serum proteins (e.g., albumin, hemoglobin, and hematocrit) in blood and of total protein in urine may be measured. The same holds in clinical studies. Blood, urine, cells, and sometimes tumors or biopsies will be collected throughout a clinical trial and assessed to determine whether the drug affects these specimens. Of interest will be whether any of these changes lead to abnormality. Knowing what is abnormal assumes you know what is normal. But how do you define what the word 'normal' means?



Surprisingly, 'normal' has no standard accepted meaning in medicine [244]. For laboratory tests, normal is expressed as the reference range. If your results are 'normal,' you are inside the reference range; if you are outside the range, you are not. Generally, the reference range is the interval within which 95% of the population lies. For example, if the reference range for fasting blood glucose (FBG) is 70-100 mg/dL, 95 out of 100 people tested should have an FBG within that range.

The exact value of these limits might be hard to find in practice. They often vary for men and women, adults and children, and the elderly, across ethnicities and disease states. Sometimes, the range is expressed in traditional units like g/dL, and sometimes in the International System of Units (SI units) like millimolar (think metric system vs. English system). The values of a test can also differ across labs because each lab uses its own equipment, staff, and assay so that laboratories will set their own reference ranges. Because of differences between laboratories, in clinical trials, a single laboratory, called a central laboratory, is often used to ensure that all samples are consistently measured.

In the first human clinical study, the SAD study, the goal is to determine the highest tolerable single dose. Volunteer safety is paramount. Further study volunteers should not be dosed if the drug is unsafe at a given dose. However, early safety signals may be difficult to identify due to the high false-positive rate of laboratory tests. With 10 laboratory tests, there is a 42% chance that at least one is outside the normal range, rising to 75% with 25 tests [245]. Suppose after drug administration, a decrease in FBG is observed. Is this a chance or a drug effect? Should the trial be stopped? If changes are observed in a single subject, sponsors will examine the entire cohort and the magnitude of the change in the volunteer. If the change is large enough, the sponsor will stop the study, or if it is seen in all volunteers, they may also consider stopping the study. Otherwise, the study may continue. Fortunately, the risk of an AE in a Phase 1 clinical trial is small. In 475 trials that enrolled over 27,000 volunteers, 15 SAEs possibly related to the study drug were observed. The median rate of mild, moderate, and severe AEs per 1000 volunteers was 1147, 46, and 0, respectively [246]. Laboratory tests pose challenges due to the number of tests and the high false-positive rate, but clinical trials remain safe for volunteers despite these challenges.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Meta-Analyses of Safety Events

Meta-analyses showed that fetuses in pregnant women taking paroxetine are at increased risk for cardiovascular malformation and malformations in general. Pregnant women should discuss with their physician whether to continue on paroxetine or switch to another medication.

Individual clinical trial data may be insufficient to detect a rare safety event. For instance, a clinical trial with 500 patients may be insufficient to detect a rare AE having an incidence of 1 in 200. But suppose multiple studies were conducted: one with 500 patients, another with 300, and another with 400, each too small to draw definitive conclusions from. Meta-analysis attempts to pool the data from all the clinical trials and draw inferences from the pooled data. To do this, a whole host of criteria must be met for the analysis to be valid: are all the trials adequate and well-controlled? Do all the studies have similar endpoints measured with the same degree of rigor? Do the studies have similar degrees of exposure and follow-up? How similar are the populations?

Even after meeting these criteria, there are still issues related to the analysis. To do a meta-analysis, you must do multiple analyses, which leads to an increased risk of finding something statistically significant when, in fact, it was not significant. This is called multiplicity. You can also get something called Simpson's paradox, where the pooled analysis reverses trends seen in individual study analyses. An FDA Guidance on meta-analyses in safety analyses has been issued and deals with these issues in much greater detail [247]. Meta-analyses have identified many risks not identified at the time of drug approval, such as the relationship between suicide and antidepressant use.

As an example, the label for paroxetine has this paragraph about its meta-analysis of congenital malformations in pregnant patients taking the drug:

> *Other studies have found varying results as to whether there was an increased risk of overall, cardiovascular or specific congenital malformations. A meta-analysis of epidemiological data over a 16-year period (1992 to 2008) on first trimester paroxetine use in pregnancy and congenital malformations included the above-noted studies in addition to others (n = 17 studies that included overall malformations and n = 14 studies that included cardiovascular malformations; n = 20 distinct studies). While subject to limitations, this meta-analysis suggested an increased occurrence of cardiovascular malformations (prevalence odds ratio [POR] 1.5; 95% confidence interval 1.2 to 1.9) and overall malformations (POR 1.2; 95% confidence interval 1.1 to 1.4) with paroxetine use during the first trimester. It was not possible in this meta-analysis to determine the extent to which the observed prevalence of cardiovascular malformations might have contributed to that of overall malformations, nor was it possible to determine whether any specific types of cardiovascular malformations might have contributed to the observed prevalence of all cardiovascular malformations.*

Detecting Rare Adverse Events in Clinical Trials

Detecting a rare AE in a clinical trial is hard. If an event is detected, is it random or a black swan event? Suppose an event happens at a rate of 1 in 1000 – that is, if 1000 patients are studied, the event will, on average, occur in 1 patient. It will take thousands of patients before the signal is large enough to state with certainty that the drug causes a rare AE. The Rule of 3 is a useful heuristic to determine how many subjects are needed to detect an AE. The Rule states that you need 3-times as many patients as the incidence to observe an AE. For example, to detect an AE with a frequency of 1 in 100, 1 in 200, or 1 in 1000, you need 300, 600, or 3000 subjects, respectively, to have a 95% chance of detecting the event. Since most drugs are developed using relatively small clinical trials, the probability of detecting an AE with a frequency of < 1 in 200 is low. For this reason, AEs of low probability are often not reported in the package insert. You will see tables of AEs that happened above a particular frequency, e.g., AEs that happened in more than 10% of study participants, such as the AE table shown in Table 14 for Lipitor (atorvastatin). Particularly serious AEs, like Stevens-Johnson syndrome, will also be reported regardless of frequency.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 14

Representative Adverse Event Table from a Package Insert

The table below lists clinical adverse reactions occurring in ≥ 2% of patients treated with any dose of LIPITOR and at an incidence greater than placebo regardless of causality (% of patients).

Adverse Event	All	10 mg	20 mg	40 mg	80 mg	Placebo
	N=8755	N=3908	N=188	N=604	N=4055	N=7311
Nasopharyngitis	8.3	12.9	5.3	7.0	4.2	8.2
Arthralgia	6.9	8.9	11.7	10.6	4.3	6.5
Diarrhea	6.8	7.3	6.4	14.1	5.2	6.3
Pain in extremity	6.0	8.5	3.7	9.3	3.1	5.9
Urinary tract infection	5.7	6.9	6.4	8.0	4.1	5.6
Dyspepsia	4.7	5.9	3.2	6.0	3.3	4.3
Nausea	4.0	3.7	3.7	7.1	3.8	3.5
Musculoskeletal pain	3.8	5.2	3.2	5.1	2.3	3.6
Muscle Spasms	3.6	4.6	4.8	5.1	2.4	3.0
Myalgia	3.5	3.6	5.9	8.4	2.7	3.1
Insomnia	3.0	2.8	1.1	5.3	2.8	2.9
Pharyngolaryngeal pain	2.3	3.9	1.6	2.8	0.7	2.1

Adverse Reactions ≥ 2% in any dose greater than placebo. Data reprinted from Lipitor package insert.

Because of the limited number of participants in clinical trials, continuous monitoring of safety data after a drug is approved is necessary to detect rare events. The FDA has several programs to monitor drugs after they are approved. The FDA Adverse Reporting System (FDAERS) is one tool that allows scientists at the FDA to use pharmacoepidemiological techniques to detect safety signals. MedWatch is another program designed for health professionals to report serious reactions or problems with approved drugs. MedWatch feeds into the FDAERS. Suppose a new AE is detected and is of particular concern; the label may need to be amended. In that case, a "Dear Doctor" letter may be mailed to all doctors notifying them of the issue, re-evaluation of the drug's approval, or all three may occur. For example, Gavreto (prealsetinib) is a drug used to treat adult patients with non-small cell lung cancer. In October 2024, the manufacturer, Rigel Pharmaceuticals, issued a Dear Doctor letter announcing a change to the Warnings and Precautions section of the label about the occurrence of severe and fatal infections while taking the drug.

Rare adverse events typically form the basis of legal issues for companies after approval. Patients who experience a serious AE after taking a drug can sue the company for negligence and product liability, even if the AE is reported on the package insert.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

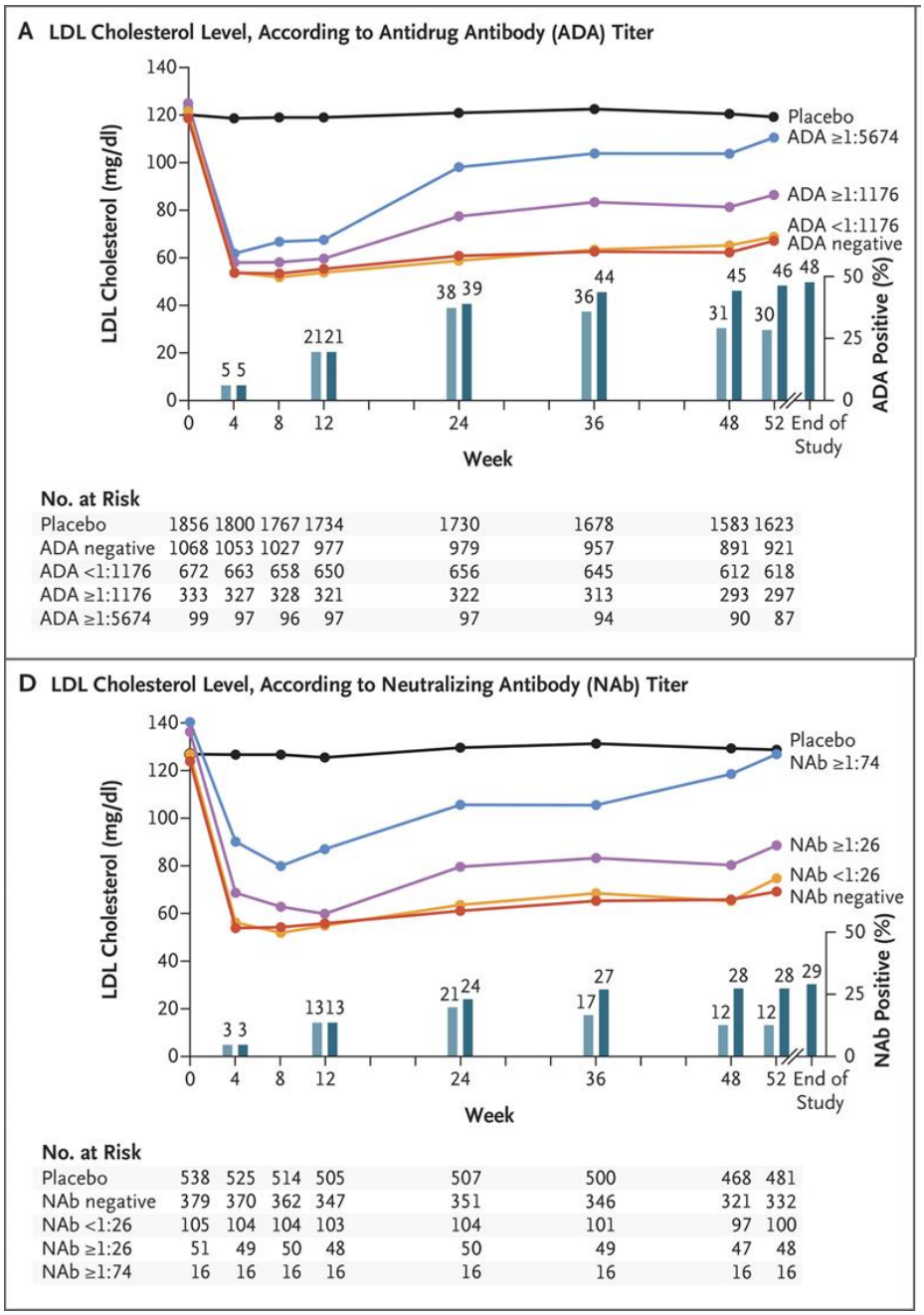


Figure 69: LDL cholesterol levels by week of treatment group by ADA titer (top) or neutralizing antibody titer (bottom) in patients treated with bococizumab or placebo. Patients with ADA titers show reduced efficacy compared to patients that are ADA-.

Reprinted with permission from [248].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

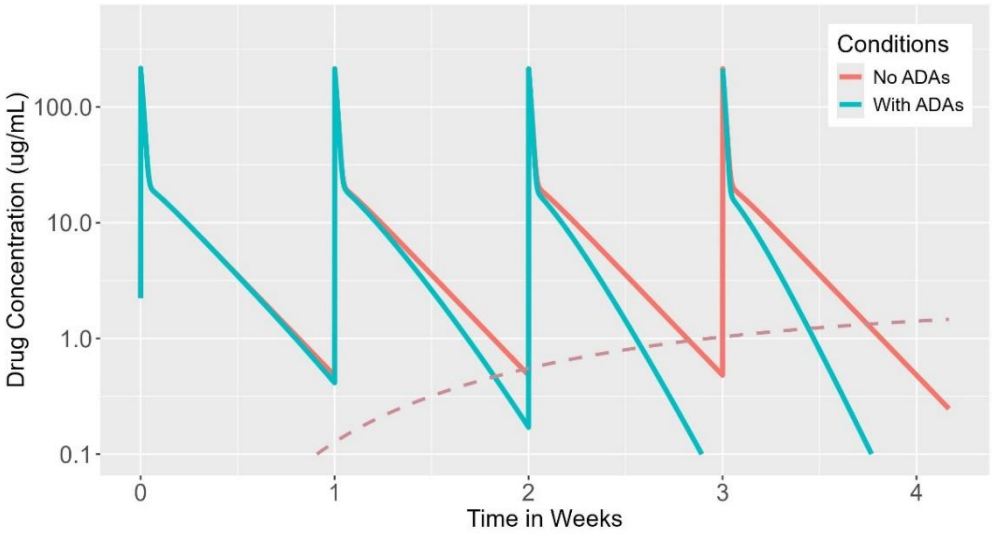


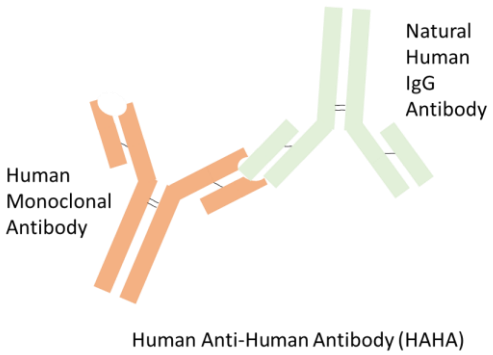
Figure 70: Drug concentration time profiles for a monoclonal antibody under two conditions: when no ADAs are present and when ADAs are present. The dashed red line shows the formation of ADAs over time. As the ADAs increase in concentration, drug concentrations are eliminated faster because the ADA-drug complex represents another route of elimination for the drug. For illustration purposes, ADA formation is faster in this example than is typically seen in humans.

Immunogenicity

Biological drugs, when administered to humans, can trigger the body to form antibodies against them. The body recognizes these as foreign materials and mounts an immune response to remove them. The antibodies that are formed are called anti-drug antibodies (ADAs). Such biologics that elicit ADAs are said to be 'immunogenic,' and when ADAs form, the patient is said to be seropositive (ADA+). Before being seropositive, the patient is seronegative or ADA-. Antibody titers measure the amount of ADA in the blood; the higher the titer, the more ADA is present. ADAs that block the binding of the target drug and inhibit the effect of the drug are said to be 'neutralizing antibodies.' By definition, neutralizing antibodies pose a greater safety concern than ADAs. Small molecules are usually not immunogenic, but can become immunogenic when bound to a carrier protein called a hapten. In such a case, the small molecule is said to be a 'hapten.' The anaphylactic reaction that sometimes happens following penicillin administration is thought to be hapten-mediated.

Although some ADAs do not affect a drug's PK, safety, or efficacy, some ADAs are of concern because they can augment or, more likely, reduce the drug's effect, increase the risk of adverse events, and possibly affect the drug's PK. Here are a couple of examples:

- ADA decreases a drug's efficacy. One of the most famous examples of how ADAs can undermine a drug's efficacy is bococizumab, a monoclonal antibody being developed to treat patients with high cholesterol. A meta-analysis of 4300 patients enrolled in 6 clinical trials showed that almost half of the patients developed ADAs after 1 year of therapy, most developing by 12



Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

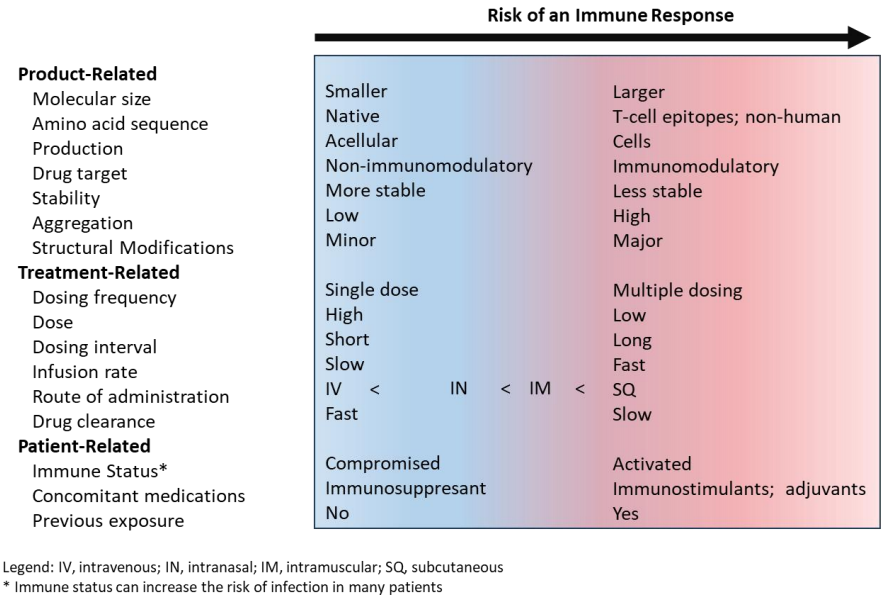


Figure 71: Factors influencing immunogenicity.

Figure redrawn from [253].

weeks. ADA+ patients did not have the magnitude of LDL reduction as ADA- patients; the magnitude correlated with ADA titers (Figure 69). Even though patients who received bococizumab for 12 weeks showed an overall 54% reduction in LDL cholesterol levels, given the wide variability in effect due to ADAs, particularly among those with the highest ADA titers, Pfizer decided to terminate the program, shocking the pharmaceutical industry [248, 249].

- ADA decreases the drug's efficacy and increases side effects: Remicade (infliximab) is a monoclonal antibody used to treat Crohn's disease and other immune-related indications. In 616 patients treated with infliximab, 24% were found to have ADAs to the drug. Disease remission at Week 30 was 35% in patients with ADAs and 54% in patients without ADAs. The rate of infusion reactions was also higher in patients with ADAs, as was disease worsening, and patients who discontinued infliximab therapy because of side-effects [250].
- ADA enhances drug effect: Some studies indicate that exogenously administered antibodies to hormones, like growth hormone and insulin, can enhance their activity, but these results were found in animals [251, 252], and it is unclear whether the same response occurs with ADAs in humans. Most current thinking is that the enhancement of a drug from ADAs is unlikely.
- ADA alters a drug's PK: ADAs can form an immune complex with a drug that results in increased elimination of the drug from the body. The addition of the complex adds another route of elimination of the drug from the body, decreasing the drug's half-life and decreasing drug concentrations in general:

$$\text{Drug} + \text{ADA} \leftrightarrow \text{Drug-ADA complex}$$
$$\downarrow \qquad\qquad \downarrow$$
$$\text{Clearance} \qquad\qquad \text{Clearance}$$

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

An illustration of this concept is illustrated in Figure 70 for a monoclonal antibody that is administered once-weekly. Over time, ADAs start to form to the drug, which causes the drug to be eliminated faster than if no ADAs were present. This is because the ADA-drug complex represents another route of elimination for the drug.

Not every biologic drug is immunogenic and not every patient will develop ADAs to the drug. Many factors can play a role (Figure 71). In a review of oncology drugs, about 2 out of every 3 biological cancer drugs are immunogenic, with a higher risk of ADA formation in drugs with non-human protein sequences, such as when the drug was mouse-derived or chimeric [254]. However, even biologics with completely human sequences still carry a risk of ADA formation. In some drugs, the incidence of ADAs can exceed 90%, while for other drugs, it can be just a few percent. Even within drugs of the same class, the incidence across drugs can vary considerably[6] [254, 255].

ADAs are considered an adverse event, and their incidence and effect are reported in the Immunogenicity subsection within the Adverse Events Section of the label. If inadequate information is available regarding ADAs, a statement to the effect of

> *"There is insufficient information to characterize the anti-drug antibody response to […] and the effects of anti-drug antibodies on pharmacokinetics, pharmacodynamics, safety, or effectiveness of […]."*

will be reported [256].

A more detailed section will be presented for drugs that have adequate information on ADAs. As an example, here is the immunogenicity section for Blincyto (blinatumomab), a cancer drug used to treat leukemia:

> *The observed incidence of anti-drug antibody is highly dependent on the sensitivity and specificity of the assay. Differences in assay methods preclude meaningful comparisons of the incidence of anti-drug antibody in the studies described below with the incidence of anti-drug antibodies in other studies, including those of BLINCYTO.*
>
> *The immunogenicity of BLINCYTO has been evaluated using either an electrochemiluminescence detection technology (ECL) or an enzyme-linked immunosorbent assay (ELISA) screening immunoassay for the detection of binding anti-blinatumomab antibodies. For patients whose sera tested positive in the screening immunoassay, an in vitro biological assay was performed to detect neutralizing antibodies.*
>
> *In clinical studies, less than 2% of patients treated with BLINCYTO tested positive for binding anti-blinatumomab antibodies. Of patients who developed anti-blinatumomab antibodies, 7 out of 9 (78%) had in vitro neutralizing activity. Anti-blinatumomab antibody formation may affect pharmacokinetics of BLINCYTO. Overall, the totality of clinical evidence supports the finding that anti-blinatumomab antibodies are not suggestive of any clinical impact on the safety or effectiveness of BLINCYTO.*

In those labels where there is adequate information about ADAs in a drug development program, the section begins with a generic paragraph on the clinical relevance of ADAs, followed by a paragraph on the methods used to assess ADAs, but few physicians and patients will know the subtle differences in ADA assessment using ECL and ELISA. This will be followed by a paragraph on the incidence in the population and an assessment of the clinical impact of the ADAs. Depending on the drug, further details on the clinical impact of ADAs can be found in the Warnings and Precautions and Clinical Trials Experience sections. For many drugs, the incidence of ADAs is sufficiently low that there are too few patients who develop ADAs to get a good assessment of their clinical impact. You will see words like

[6] It should be noted that because of the differences in analytical methods, cut-off values, when the samples were collected in the clinical trial, and differences in statistical methods, it is inappropriate to compare the incidence of ADAs across drugs. For instance, it would not be appropriate to say that Drug X has lower immunogenicity risk because the incidence of ADAs is 10% compared to 20% for Drug B.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

"Every medicine is a balance between its intended benefits and its potential risks; the goal of science and regulation is not to eliminate all risk, but to ensure that the benefits justify it." – Janet Woodcock

'are not suggestive' or something like "*Because of the low occurrence of anti-drug antibodies, the effect of these antibodies on the pharmacokinetics, pharmacodynamics, safety, and/or effectiveness of polatuzumab products is unknown*," which is how Polivy (polatuzumab vedotin-piiq) reports their immunogenicity results.

Benefit-Risk Assessment

For a drug to be approved, it must show both efficacy *and* safety. Since all drugs have adverse events, it must be shown that the benefits of taking the drug outweigh the risks of taking the drug. Recall many chapters ago, the drug Seldane (terfenadine). Seldane was approved for the treatment of allergic rhinitis, a condition characterized by a runny nose. But in certain patients, the drug could cause fatal arrhythmias. In this case, it was determined that the benefits of taking the drug did not outweigh its risks, and the drug was eventually removed from the market. The FDA makes a thorough, case-specific, benefit-risk assessment for all new drug applications. This assessment reviews the efficacy, safety, and any data gaps [257].

Sometimes, the assessment is straightforward, such as for a drug that prolongs survival in life-threatening diseases without serious adverse events. However, in other cases, it can be murky, such as when the potential for serious adverse events exists. In such cases, the FDA may consider whether measures or companion diagnostics are available to mitigate the risk of the adverse event, or whether a lower dose can be used that still achieves similar efficacy. They will also examine whether other approved drugs are indicated for this indication, or whether this is the first drug to fill an unmet medical need. A greater risk may be accepted if there are no alternative therapies. Are there subpopulations that are more at risk than others? If so, a greater warning in the package insert can be made regarding dosing in that population. The FDA may also consider the benefit-risk outside the population, such as with a drug that has abuse potential and may find its way into individuals who do not have the disease of interest. These considerations can sometimes lead to conflict because the FDA is concerned with the benefit-risk in a broader health perspective rather than in a particular individual. The FDA may deny approval of a drug because of its overall poor benefit-to-risk profile, even though there is evidence that the drug works and is safe for specific individuals who have taken it.

The vehicle the FDA uses to assess and communicate a drug's benefit-risk is its Benefit-Risk Framework. This structured, qualitative approach identifies the important factors in making the assessment. The major factors examined include an analysis of the condition, current treatment options, the drug's benefits, and benefit-risk management. These factors are examined in terms of evidence, uncertainty, conclusions, and reasons. An Overall Benefit-Risk Conclusion is made at the end. As a result of the analysis,

The safety and effectiveness of a drug depend on its context of use. A drug may show a favorable risk-benefit ratio for one indication, but may be less so for another. A drug may possess a favorable risk-benefit ratio for a severe, terminal condition but a poor one for a minor illness.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

the FDA may include black box warnings on the label, require additional postmarketing studies and commitments, or require the sponsor to provide risk evaluation and mitigation strategies (REMS), a safety program enforced by the FDA to help mitigate drug risks. An example of REMS can be found in Zyprexa Relprevv, an antipsychotic whose use can lead to a post-injection delirium sedation syndrome whose symptoms include sedation, coma, and feeling confused or disoriented within 1 to 3 hours after administration. Although the overall incidence is < 1%, the sponsor was required to implement a REMS plan that requires the drug to be administered only in health care facilities that can monitor the patient for at least 3 hours and provide medical care if needed [258]. The REMS requirement is communicated on the package insert by a black box warning:

> ***WARNING: POST-INJECTION DELIRIUM/SEDATION SYNDROME and INCREASED MORTALITY IN ELDERLY PATIENTS WITH DEMENTIA-RELATED PSYCHOSIS***
>
> *See full prescribing information for complete boxed warning.*
>
> *Patients are at risk for severe sedation (including coma) and/or delirium after each injection and must be observed for at least 3 hours in a registered facility with ready access to emergency response services. Because of this risk, ZYPREXA RELPREVV is available only through a restricted distribution program called ZYPREXA RELPREVV Patient Care Program and requires prescriber, healthcare facility, patient, and pharmacy enrollment. (2.1, 5.1, 5.2, 10.2, 17)*
>
> *Elderly patients with dementia-related psychosis treated with antipsychotic drugs are at an increased risk of death. ZYPREXA RELPREVV is not approved for the treatment of patients with dementia-related psychosis. (5.3, 8.5, 17)*

Assessment of the drug's benefit-risk does not end with the drug's approval. The FDA reviews the drug's benefit-risk after marketing authorization, recognizing that not all the risks are known at the time of approval. FDA reviews medical literature, the company's annual safety reports, patient experience data, product quality reports, medication error reports, and other sources to inform the drug's risk profile. The FDA further requires sponsors to regularly and periodically re-evaluate the benefit-risk profile of their drugs. Sponsors must "*[P]resent a comprehensive, concise, and critical analysis of new or emerging information on the risks of the medicinal product and on its benefit in approved indications, to enable an appraisal of the product's overall benefit-risk profile*". If serious concerns arise, sponsors are expected to inform the FDA immediately. Further information can be found in the FDA's Guidance to Industry on Benefit-Risk Assessment [257].

Not All Adverse Events Are Bad

Not all AEs are bad. There are cases where, because of reported AEs with a new drug, the sponsor changed the direction of the drug development program and took the drug into a new indication. When scientists at Pfizer first studied sildenafil, they thought it might be useful for treating hypertension. When the nurses checked on the male volunteers, they found them lying on their stomachs. Upon further querying, they noticed an unusual side effect – the men had erections and were lying on their stomachs to hide it from the nurses. Upon further testing, using more sex-related endpoints, the drug proved its efficacy, and in 1998, Viagra was approved for the treatment of erectile dysfunction. The world has not been the same ever since.

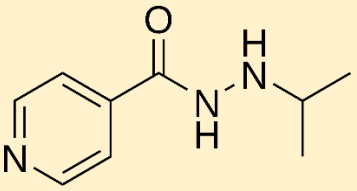
Minoxidil was first tested as an antihypertensive, but the men taking it began to grow unwanted hair. Because of this, Upjohn began testing the medication as a hair loss treatment and found that topical administration could prevent hair loss or regrow lost hair in some men. In 1979, minoxidil was approved as an antihypertensive drug, and in 1988, it was approved for hair loss in men as Rogaine (the FDA rejected the original name of Regain).

Isoniazid was developed as a treatment for tuberculosis, but in 1952, researchers noticed that some

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

patients were inappropriately happier after taking the medication. Subsequently, after chemically modifying isoniazid to improve absorption, the resulting drug, iproniazid, was sold as an antidepressant for almost a decade before it was removed from the market because of liver toxicity. Nevertheless, iproniazid's similarity to natural neurotransmitters in the brain led psychiatric researchers to study further how modification of natural neurotransmitters could lead to other antidepressants.

AEs can also be a sign a drug may have a positive therapeutic outcome. Many old cancer drugs, such as cytotoxic drugs, cause white blood cell populations to decline. For some cancers, severe neutropenia is correlated with greater survival, like in myelodysplastic syndrome. Even for targeted therapies, which are generally safer than cytotoxic drugs, some drugs, like Tarceva (erlotinib), can result in a profound rash. In non-small cell cancer patients treated with Tarceva, a 2-month improvement in survival was observed among those who developed the rash. The overall response rate to the drug was 17% in rash patients compared to 3% in patients without rash [259]. The 'Tarceva Rash' is seen as a sign that the patient will be more likely to respond to the drug.



Viagra, Rogaine, and Iproniazid were all discovered by having unusual ADRs

Summary

Today, we expect our drugs to be perfectly safe. They are not. No drug is completely safe. Paracelsus, the Father of Toxicology, said almost 500 years ago that all drugs are poison; it's just a matter of dose. When taking a drug, it's a matter of balancing the benefits of the drug with the side effects of the drug and maintaining the benefit-to-risk ratio. For certain classes of drugs, a higher side-effect profile is accepted. For example, you might be willing to accept a drop in white blood cell count and an increased risk of infection to cure your cancer, but you would not accept such a risk for a drug to treat a runny nose. If you experience an AE with a drug, you should report it to your physician. Alternative medications or dosing strategies may also be available.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 8

Special Populations

Introduction

The Use in Specific Populations section of a drug's label, previously known as the Special Populations section, describes those patient subgroups that may require dose modifications or increased physician oversight due to safety or efficacy concerns. The subgroups within the Specific Populations section may have altered pharmacokinetics or altered pharmacodynamics. This section may include specific information related to pregnant or lactating patients, patients with renal or hepatic impairment, children, and, possibly, the elderly. The Specific Populations section may also include information on patient subgroups that are underrepresented in clinical trials, either because of the inclusion/exclusion criteria used in pivotal trials, the difficulty of studying these subgroups, or simply because few patients from these subgroups were enrolled.

Not all patients are eligible to enroll in a specific clinical trial. Even if you have diabetes, for example, you might not qualify for enrollment in a clinical trial to test the safety and efficacy of a new drug to treat diabetes. All clinical studies have inclusion and exclusion criteria for enrollment. A patient must meet all of the inclusion criteria to be considered for enrollment. Meeting any exclusion criterion disqualifies a patient from the study. A patient may enroll in a clinical trial only if they meet all of the inclusion criteria and none of the exclusion criteria.

Adding inclusion/exclusion criteria (I/EC) to a clinical study has advantages and disadvantages. I/EC ensures a homogeneous population, focuses the study towards the desired target population, and controls for variables that may influence the study results, thereby increasing the probability of a successful trial. Researchers have an ethical obligation to ensure that trial participants are not harmed. Exclusion criteria can be used to minimize these risks by excluding participants who are at increased risk or especially vulnerable. By doing so, however, if the researcher is not careful, inappropriate exclusion criteria can introduce unintended bias into the study results and limit the generalizability of the findings to the broader population. Further, it is also important that no exclusion criterion affects the study results. Hence, a delicate balance must be struck between studying a homogeneous population and still allowing the results to generalize to the real-world population.

An example of the inclusion/exclusion criteria used in a clinical trial is Mounjaro (tirzepatide), which was approved as an adjunct to diet and exercise to improve glycemic control in patients with Type 2 diabetes. To qualify for the Surpass-1 trial, the pivotal Phase 3 clinical trial used to support the approval of Mounjaro, the following inclusion criteria were used:

- Diagnosed with type 2 diabetes mellitus.
- Naïve to diabetes injectable therapies and have not used oral antihyperglycemic medications during the 3 months preceding screening.
- Hemoglobin A1c between ≥7.0% and ≤9.5%.
- Stable weight (± 5%) for at least 3 months before screening.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026.

- Body mass index (BMI) ≥23 kg/m^2 at screening.

Patients must have met all these criteria to qualify for the trial. If, for example, a patient had a hemoglobin A1c of 10% (normal hemoglobin A1c is < 5%), they would not have qualified for the trial because their hemoglobin A1c was outside the acceptable range for enrollment. Furthermore, if a patient had a BMI of 22 kg/m^2 (normal BMI in the United States is 19 to 25 kg/m^2), which is not underweight, they would also not qualify for the study because patients were required to have a BMI greater than 23 kg/m^2. The exclusion criteria for Surpass-1 were:

- Have type 1 diabetes mellitus.
- Have had chronic or acute pancreatitis at any time before study entry.
- Have proliferative diabetic retinopathy, diabetic maculopathy, or nonproliferative diabetic retinopathy requiring acute treatment.
- Have disorders associated with slowed emptying of the stomach or have had any stomach surgeries for weight loss.
- Have an estimated glomerular filtration rate <30 mL/minute/1.73 m².
- Have had a heart attack, stroke, or hospitalization for congestive heart failure in the past 2 months.
- Have a personal or family history of medullary thyroid carcinoma or a personal history of multiple endocrine neoplasia syndrome type 2.
- Have been taking weight loss drugs, including over-the-counter medications during the last 3 months.

Meeting any of these criteria would disqualify a patient from enrollment. If a patient had been taking Alli (orlistat), an over-the-counter weight loss drug, 2 months before screening for the study, this patient would be disqualified from enrolling in the study, or if 20 years ago, the patient had gallstones because they were a heavy drinker (but have been sober since) and had a bout of acute pancreatitis, they would also be disqualified from the study.

Due to the inclusion/exclusion criteria, certain population segments may be underrepresented in clinical trials. As an example, Veozah (fezolinetant) was approved for the treatment of moderate to severe vasomotor symptoms, i.e., hot flashes, due to menopause. One of the pivotal Phase 3 clinical trials supporting development was Skylight-1. To qualify for the study, the subject must have had a BMI of ≥18 kg/m2 and ≤38 kg/m^2. Morbidly obese is a BMI of 40 kg/m^2 or higher. Do morbidly obese patients get hot flashes? Yes, higher body weight is associated with a higher frequency of hot flashes. But, using the study's inclusion criteria, an inadvertent special group has been created. It is well established that obesity can impact the pharmacokinetics and pharmacodynamics of certain drugs (see later in this chapter). By excluding morbidly obese patients from the study, it is unknown whether these patients need any dose adjustment in clinical practice; the label is silent on this issue.

The population used to study a drug in Phase 2 or Phase 3 trials is expected to reflect the population intended for its use [260]. Most clinical studies seek to exclude participants with co-morbidities; they rarely accept all-comers with the disease. Drug approval is based on studying large, relatively homogenous groups in prospective clinical trials and extrapolating the results to anyone prescribed the drug. It is hard to find a consistent, reasonable definition of what constitutes a special population. Still, a simple one is any subgroup that may be of special interest to physicians.

Referring back to the question regarding obesity and Veozah, what is the recommended dose for an obese patient? Suppose you are a physician and an obese patient who experiences hot flashes comes to you and asks to be put on Veozah after seeing one of the television commercials for the product. What do you do? Do you tell them no? Would the drug even work for that patient? After all, obese patients were not studied in the pivotal clinical trials. The word 'obese' isn't even mentioned in the drug label. What dose should the patient take? Should they take the recommended dose, 45 mg once-daily, or should they take a larger one because they are a bigger person? And if the patient were to take a higher dose, what would that dose be? Double the dose? Or maybe 45 mg twice a day instead of once a day? These are the problems posed by special populations. How do you dose them? The concept of "one dose fits all", even though that is how drug companies try to develop and market

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

their drugs, does not apply in every case. We are all different, and different patients may require a different dose than others.

Today, the specific populations usually of interest to physicians include pregnant patients, lactating patients, elderly patients, pediatric patients, obese patients, patients with renal and/or hepatic impairment, patients with genetic differences, or patients with rare diseases. What constitutes a special population today may not do so in the future. For example, before the 1990s, women were considered a special population. Drug manufacturers were worried that the inclusion of women in clinical studies might lead to an unacceptable toxicity profile for the drug, or what would happen if a woman became pregnant while taking the drug.

A drug with a clinical consequence of this attitude is zolpidem (Ambien), a frequently prescribed sleeping aid that was approved in 1992. After millions of prescriptions, it was found that women had a significantly higher number of adverse events compared to males, and further studies show that this was due to women having higher blood concentrations than males at the same dose, even after accounting for weight. Zolpidem was being prescribed at the same dose for males and females. Today, the standard zolpidem dose for women is 5 mg compared to 10 mg for men. In 1993, the FDA started requiring greater female participation in clinical trials. Today, men and women are expected to be equally allowed to participate in clinical trials unless the drug is being specifically developed for a particular sex. For instance, hormone replacement therapy is usually only studied in males or females, depending on which hormone is being studied.

Information related to special populations can usually be found in Section 8 of the drug label, titled "Use in Specific Populations." The information provided in this section will help physicians prescribe the drug to these patients. For example, baricitinib (Olumiant) was approved in 2018 for the treatment of rheumatoid arthritis. In Section 8.7, Renal Impairment, the drug label states:

> *Renal function was found to significantly affect baricitinib exposure. OLUMIANT is not recommended for use in patients with estimated GFR of less than 60 mL/min/1.73 m^2 [see Dosage and Administration (2.4) and Clinical Pharmacology (12.3)].*

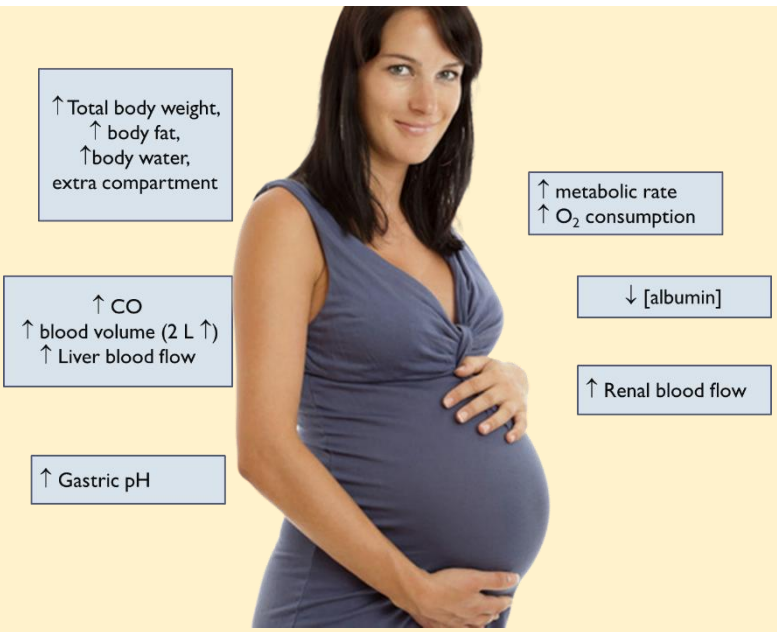
Hence, baricitinib is not recommended in patients with moderate to severe renal impairment.

The rest of this chapter will focus on several special populations: pregnancy and lactation, renal and hepatic impairment, the elderly, pediatric, and obesity. Some of the physiological changes that occur in these subgroups and how they might affect dose requirements in these subgroups will be discussed.

Pregnancy and Lactation

To say that physiological changes occur during pregnancy would be an understatement, but what might not be so obvious is that pregnancy can affect the safety and efficacy of a drug through modulation of the drug's pharmacokinetics and/or pharmacodynamics. Pregnant women are usually excluded from participating in clinical trials because of safety concerns and, as such, represent a special population that may require a different dosing regimen than non-pregnant patients. While it is understandable that drug manufacturers avoid studying their drugs in pregnant women for safety reasons, after all, what manufacturer wants another thalidomide scandal on their hands? What is surprising is that regulatory authorities have not yet required studies in pregnant women for any new drug that may be prescribed to patients who become pregnant, considering how many women become pregnant every year and the fact that about 1 in 20 women are pregnant at any given time. While many women stop taking medication while they are pregnant, there are many drugs for which they cannot. It's been estimated that during the first trimester, roughly 70 to 80% of pregnant women take medication of some kind, and that the incidence of taking medications during pregnancy is increasing over time [261, 262]. Many physiological changes occur during pregnancy at different times [263]. There is an obvious change in body weight, the addition of the fetal compartment, an increase in body fat and water, but in addition, there are changes in metabolic rate, blood flow to organs, and changes in gastric pH, all of which can affect a drug's pharmacokinetics. Every process

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Many physiological changes occur during pregnancy, which may affect a drug's pharmacokinetics.

in a drug's pharmacokinetics can be affected: absorption, distribution, metabolism, and elimination.

Consider midazolam and digoxin as examples (Figure 72). Researchers at the University of Washington administered midazolam and digoxin to women 28 to 32 weeks into their pregnancy and 10 weeks postpartum. Midazolam is an excellent probe for CYP3A metabolism and changes in its pharmacokinetics reflect changes in CYP3A metabolism. Digoxin, on the other hand, is an excellent probe for studying the activity of p-glycoprotein (P-gp). Recall that p-glycoprotein is a protein transporter that sits in the cell membrane and pumps drugs out of cells, and acts to lower drug concentrations in the blood. Changes in digoxin concentrations reflect changes in absorption and renal elimination. Both midazolam and digoxin showed decreased concentrations during pregnancy. This is a common observation: pregnancy tends to increase enzyme activity because overall metabolism increases to support the pregnancy. Hence, enzymes, like CYP3A4, and transporters, like p-glycoprotein, have increased activity, both of which will act to decrease drug concentrations.

The impact of these physiological changes depends on the pregnancy trimester. Figure 73 shows simulated tacrolimus blood concentrations during pregnancy. Tacrolimus is a drug used to prevent organ rejection after receiving a transplant. Tacrolimus requires blood concentrations within a specific range, typically 5-15 ng/mL, for it to be effective. Blood concentrations in pregnant patients decrease each trimester, reaching their nadir in the third trimester, where concentrations are 25% to 40% lower than pre-pregnancy levels. Even 10 weeks postpartum, tacrolimus concentrations remain lower than before pregnancy. Because of its narrow therapeutic window, careful monitoring of tacrolimus drug concentrations during pregnancy is required.

"We know shocking little about medication safety during pregnancy" –

Rachel Feltman,
Scientific American [1]

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

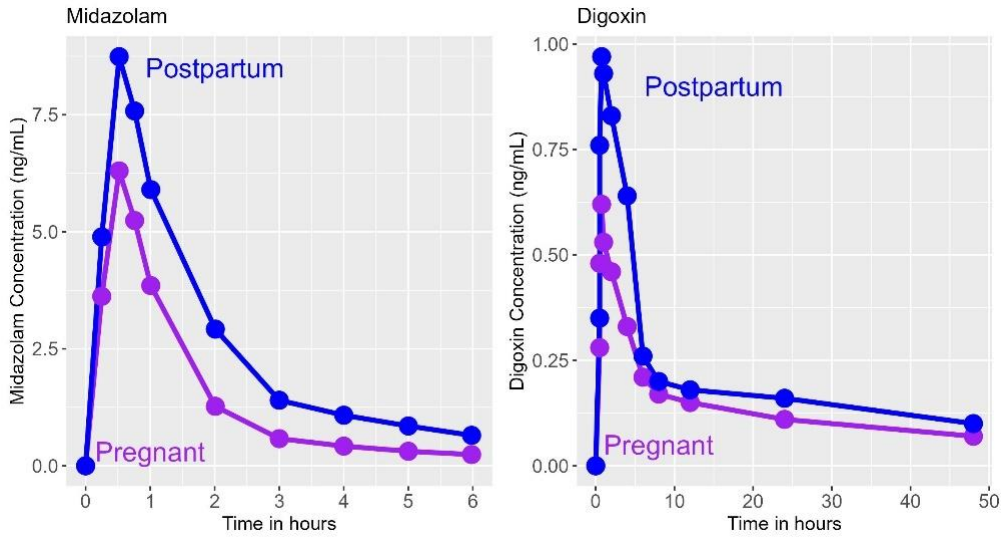


Figure 72: Midazolam and digoxin pharmacokinetics during pregnancy and postpartum 10 weeks after pregnancy. Drug concentrations of both drugs are reduced during pregnancy. Figures redrawn from [264].

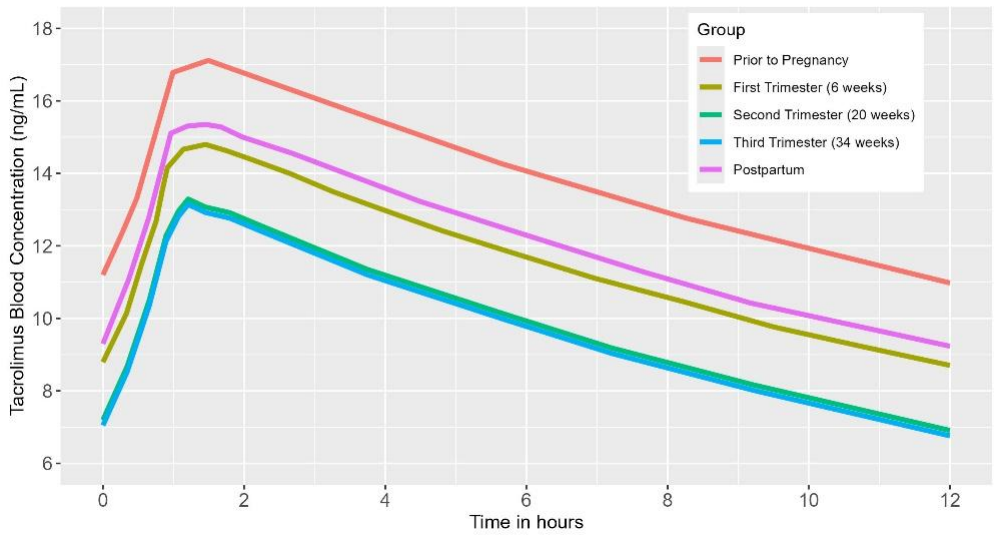


Figure 73: Tacrolimus drug concentrations change in different trimesters during pregnancy. Values were simulated based on population analysis of tacrolimus blood concentrations from 14 pregnant women with a prior kidney transplant. Notice that the 2nd and 3rd trimesters overlap. The therapeutic window for tacrolimus is 5 to 15 ng/mL. During pregnancy, tacrolimus blood concentrations are lower, potentially requiring dose adjustment. Even immediately postpartum, tacrolimus concentrations remain lower than when not pregnant. Figure redrawn from [265].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Drugs in Breast Milk

Drugs ingested by a lactating mother may be transferred to a breastfeeding infant and are cause for concern. During lactation, the breast increases the number of lactiferous ducts and the size and number of breast lobules (which produce milk), allowing maternal nutrients to be transferred from maternal blood into breast milk.

Passive diffusion is the primary route of drug transfer into breast milk. Knowing what we know about how drugs diffuse across a cell membrane, we can predict the characteristics a drug will need to cross from plasma to breast milk. [266]: unionized drugs transfer more readily than ionized drugs, lipophilic drugs transfer more readily than polar drugs, small molecules transfer more readily than biologics, drugs with high plasma concentrations in the mother transfer to a greater extent than drugs with low plasma concentrations, and drugs with low binding to plasma proteins in the mother are more likely to transfer than drugs with high protein binding.

Infant-related factors also play a role. How much drug gets absorbed by the infant contributes, as does the infant's ability to metabolize the drug during first pass through the liver. Timing of breastfeeding relative to the time of drug administration in the mother may also play a role, as does how much milk the baby consumes.

Some drugs are actively secreted from plasma to milk by transporter proteins, which can result in drug concentrations much higher in breast milk than in the mother's plasma. For example, doxorubicin, an anticancer drug, has a 10-fold greater concentration in milk than in plasma. Some transporters secrete drugs from milk back into plasma, although this is rare. Nifedipine, a heart medication, is one such example, and it has 25% lower concentrations in milk than in plasma due to active secretion back into the plasma.

The National Library of Medicine has a useful database called Drugs and Lactation Database that can be accessed should you have particular questions regarding a drug and whether it can get into breastmilk.

Frequently, drug companies, rather than conducting clinical trials in humans, may conduct studies in pregnant animals to examine how pharmacokinetics may change during pregnancy and whether the drug may enter breast milk, which breastfeeding infants could then absorb. For example, the text from Section 8.1, the Risk Summary, for Olumiant (baricitinib) is:

> *The limited human data on the use of OLUMIANT in pregnant women is not sufficient to inform a drug-associated risk for major birth defects or miscarriage. In animal embryo-fetal development studies, oral baricitinib administration to pregnant rats and rabbits at exposures equal to and greater than approximately 20 and 84 times the maximum recommended human dose (MRHD), respectively, resulted in reduced fetal body weights, increased embryolethality (rabbits only), and dose-related increases in skeletal malformations. No developmental toxicity was observed in pregnant rats and rabbits treated with oral baricitinib during organogenesis at approximately 5 and 13 times the exposure at the MRHD, respectively. In a pre-and post-natal development study in pregnant female rats, oral baricitinib administration at exposures approximately 43 times the MRHD resulted in reduction in pup viability (increased incidence of stillborn pups and early neonatal deaths), decreased fetal birth weight, reduced fetal body weight gain, decreased cytotoxic T cells on post-natal day (PND) 35 with evidence of recovery by PND 65, and developmental delays*

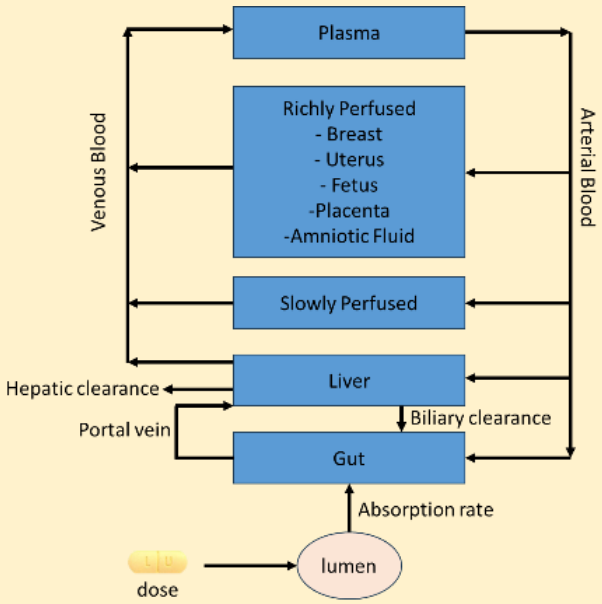
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

that might be attributable to decreased body weight gain. No developmental toxicity was observed at an exposure approximately 9 times the exposure at the MRHD [see Animal Data].

The estimated background risk of major birth defects and miscarriage for the indicated population(s) is unknown. All pregnancies have a background risk of birth defect, loss, or other adverse outcomes. In the U.S. general population, the estimated background risk of major birth defects and miscarriage in clinically recognized pregnancies is 2-4% and 15- 20%, respectively.

There is no information from humans in the drug label. The entire pregnancy section is based on changes in animals. The information allows the physician to understand the drug's benefit-risk profile in animals and extrapolate those results to humans. Furthermore, there is no information on how to dose the drug in humans.

The Minimal PBPK Model for Pregnancy Used by George et al. [267] to Simulate Sertraline PK



A recent development uses PBPK modeling to simulate changes during pregnancy and predict drug concentrations in the placenta and fetus. George et al. [267] used a minimal PBPK model, with fewer compartments than a typical PBPK model, to predict sertraline pharmacokinetics during pregnancy. Their model accounted for changes in blood flow, liver enzyme contribution during pregnancy, cardiac output, organ volumes, and patient weight during pregnancy. The model predicted a 1.4-fold increase in sertraline clearance at 40 weeks, meaning that pregnant women would need to receive almost twice their usual dose during the third trimester to achieve concentrations similar to those before pregnancy. While this is currently a hot topic of research within the pharmacometrics community, no drug label has, as of 2025, incorporated PBPK predictions to help guide dosing.

Renal and Hepatic Impairment

Liver impairment can be caused by many factors: hepatitis C, physical liver injury, acetaminophen overdose, bile flow impairment, and alcoholism. Renal impairment can also be caused by many mechanisms: hypertension, aging, hypoxia, and iron overload, to name a few. Both renal and hepatic impairment can decrease the elimination of drugs from the body. Patients with renal or hepatic impairment may require lower doses because of increased and prolonged drug concentrations in these patients. As such, sponsors must assess for dose modifications in these special populations.

Special populations often have physiological differences that may require dose adjustments. Today's standard development paradigm is not to conduct large-scale safety and efficacy trials in these special populations, but rather to conduct small, focused studies instead. To determine the dose to be used in a particular special population, drug companies first assess whether the pharmacodynamics in that population are the same as in the general population. If that is the case, the 'principle of therapeutic

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

equivalency' may be used. This principle states that if the pharmacodynamics of a drug are the same in two populations and the concentration-time profiles are identical, then the drug's safety and efficacy will be the same in both populations. As you will recall, this is the entire basis for generic drugs; it is not necessary to do a safety and efficacy phase 3 study with a generic drug; it is sufficient to perform a bioequivalence study to show that the rate and extent of exposure are the same for the generic drug compared to the reference formulation. The goal is to determine the dose in the special population that matches the exposure level in the general population from the Phase 3 study.

There are two general ways to measure renal function, both related but not identical. One way is to estimate creatinine clearance (eCrCL). Creatinine is a byproduct of muscle metabolism that is primarily eliminated through the kidneys. Creatinine clearance measures the rate at which the kidneys remove creatinine from the blood and is used to estimate the glomerular filtration rate (GFR). Recall that the glomerulus is the filtration part of the nephron and the first step in urine formation. Glomerular function is a good measure of overall kidney function. Normal values vary by the source but are approximately greater than 90 mL/min/1.73 m^2. Lower values indicate varying degrees of renal impairment. Generally, a rate of 60 to 90 mL/min/1.73 m^2 is considered mild impairment, 30 to 60 mL/min/1.73 m^2 is considered moderate impairment, and below 30 mL/min/1.73 m^2 is considered severe renal impairment, which may require dialysis. eCrCL is usually estimated by the Cockroft-Gault (C-G) equation [268], which is based on a person's sex, weight, age, and serum creatinine concentration (S_{Cr}):

$$\text{eCrCL } (\text{mL/min}) = \frac{(140 - \text{age in yrs})(\text{Weight in kg})}{72 \times (\text{S}_{cr} \text{ in mg/dL})} \times (0.85 \text{ if female}) \tag{34}$$

eCrCL is not a perfect measure of glomerular function, because some of the creatinine that passes through the glomerulus is reabsorbed from the nephron back into the blood. There are also issues regarding eCrCL estimation with obese or underweight individuals.

Today, eCrCL is not recommended as a measure of renal function and, instead, estimated GFR (eGFR) is preferred. Two equations for estimating GFR are commonly used. The Modification of Disease in Renal Disease (MDRD) has two versions, one having six predictor variables and the other having four predictor variables [269]. The four-predictor equation is more commonly used:

$$\text{eGFR } \left(\text{mL/min/1.73 m}^2\right) = 175 \times \text{S}_{\text{cr}}{}^{-1.154} \times \text{age}^{-0.203} \times (1.212 \text{ if African American})(0.742 \text{ if female}) \tag{35}$$

MDRD has been criticized because of its use of race in the model and its overestimation of eGFR in African Americans. Current guidelines recommend using the CKD-EPI equation [270]:

$$\text{eGFR } \left(\text{mL/min/1.73 m}^2\right) = 142 \times \min\left(\frac{\text{S}_{\text{cr}}}{\kappa}, 1\right)^{\alpha} \times \max\left(\frac{\text{S}_{\text{cr}}}{\kappa}, 1\right)^{1.2} \times 0.9938^{\text{Age}} \times (1.012 \text{ if female}) \quad . \tag{36}$$

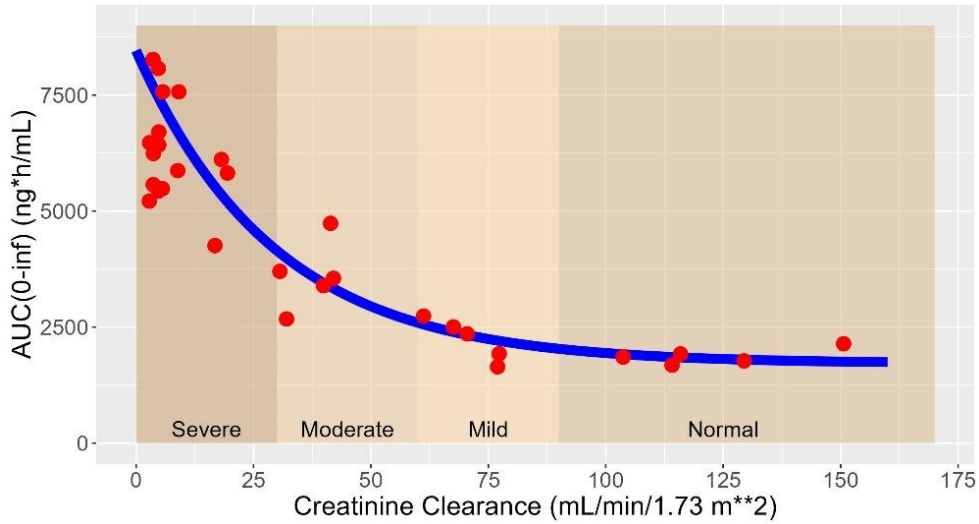
$$\kappa = 0.7 \text{ for females, } 0.9 \text{ for males}$$

$$\alpha = -0.241 \text{ for females, } -0.302 \text{ for males}$$

The downside of the CKD-EPI equation is its complexity, which almost always necessitates using a calculator or app. Care needs to be taken in comparing estimates using different equations because it is possible for patients to change groups, like from moderate to mile impairment when changing from eCrCL to eGFR.

Zerit (stavudine), an early antiretroviral drug used to treat HIV that is no longer in use, will be used as an example for how dose modifications in patients with renal impairment are estimated because stavudine is cleared predominantly through the kidneys. Figure 74 shows the AUC for stavudine in 31 volunteers with varying degrees of renal impairment plotted against their creatinine clearance.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Stavudine Dosing Recommendations for Patients with Impaired Renal Function		
CrCL (mL/min/1.73 m^2)	Recommended dose for body weight of:	
	≤ 60 kg	> 60 kg
> 50	40 mg every 12 h	30 mg every 12 h
21 – 50	20 mg every 12 h	15 mg every 12 h
10 – 20	20 mg every 24 h	15 mg every 24 h
Hemodialysis	20 mg every 24 h	15 mg every 24 h

Figure 74: Stavudine AUC as a function of creatinine clearance (top) following single dose administration of 40 mg stavudine in patients with varying degrees of renal impairment and corresponding dosing table. Figure redrawn from [271].

The figure shows that volunteers with mild renal impairment and normal renal function tend to have AUC values around 1500 ng*h/mL, volunteers with moderate renal impairment tend to have AUC values around 3500 ng*h/mL, and those with severe renal impairment have AUC values around 6000 ng*h/mL, 4-fold higher than volunteers with normal or mild renal impairment. Increased AUC means higher stavudine concentrations, which may lead to increased toxicity.

Intuitively, you want concentrations in patient subgroups to match those in patients with normal renal function in the general population. In this case, if you want to match concentrations in patients with severe renal impairment with those in normal patients, you will need to decrease the stavudine dose by 75%, from 40 mg to 10 mg. This is what the company did. The company 'exposure-matched' renally impaired patients and created a dosing table that produces approximately the same AUC as in patients with normal renal function, which they placed in the drug label. Since stavudine was administered at different dosages based on body weight, the dosing regimen was also divided into two groups: those above and below 60 kg. The dosing table for stavudine in patients with impaired renal function is also shown in Figure 74. A 10 mg tablet was unavailable, but a 20 mg tablet was. Instead of reducing the dose from 40 to 10 mg, the sponsor reduced it to 20 mg and increased the dosing interval from twice-daily to once-daily, achieving the same result. This example elegantly

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

illustrates how sponsors create dosing regimens for patients with impaired renal function. A similar approach is used for patients with hepatic impairment.

The experimental design the authors used to assess the effect of renal impairment on stavudine pharmacokinetics is referred to as a full design. Six to eight volunteers from each impairment group (normal, mild, moderate, severe) are enrolled, and the drug's pharmacokinetics are assessed after a single dose. This is referred to as a full design because volunteers are recruited from each group. More typically today, companies use a reduced design in which they enroll volunteers in the severe group, and then enroll volunteers with normal renal function who are similar in sex, age, and weight. The drug's pharmacokinetics are characterized after a single dose. If there is no difference, the study is stopped and no renal impairment effect is declared, assuming that if the severe group shows no change in pharmacokinetics, the mild and moderate groups should not either. If there is a difference, the company continues to enroll the middle two cohorts and completes the design.

Even with using a reduced design, renal and hepatic impairment studies are hard to do. It's often difficult to find patients with the disease of interest who have renal or hepatic impairment. For this reason, companies typically do not enroll patients with the disease of interest but instead enroll any volunteer with the desired degree of impairment. Even with this simplification, it's just generally hard to find any volunteers with renal or hepatic impairment, particularly those with severe impairment, who are willing to participate as volunteers in a clinical trial. Additionally, for certain classes of drugs, like cancer drugs, it may be unethical to enroll otherwise healthy volunteers and administer the drug to them. For these reasons, companies are seeking alternative methods to address dose modifications in patients with renal or hepatic impairment.

An alternative approach to conducting a clinical study in volunteers with renal or hepatic impairment is to enroll these patients in Phase 2 or 3 clinical trials and then utilize population pharmacokinetics (PopPK) to assess the impact of impairment on the drug's pharmacokinetics. As you may recall, PopPK is a complex mathematical and statistical approach to characterize a drug's pharmacokinetics when there are limited samples per patient. This makes it ideal for characterizing a drug's pharmacokinetics in late-stage clinical trials where collecting multiple samples per patient is difficult. The effects of patient characteristics, such as age, weight, or renal function, can be incorporated into the models to assess their impact on drug concentrations. An example of this approach is the label for enfortumab vedotin (Padcev), an antibody-drug conjugate used to treat patients with urothelial cancer. PopPK was used to characterize the pharmacokinetics of enfortumab vedotin and its metabolite MMAE in 748 patients enrolled across multiple clinical studies [272]. Of the 748 patients, 613 had some degree of renal impairment, and of those, 25 patients had severe renal impairment. Also, 68 patients had some degree of hepatic impairment, 65 of which were mild in nature; no severe patients were enrolled. No significant effect of renal or mild hepatic impairment was observed, and this was reflected in the label:

> *"No clinically significant differences in the pharmacokinetics of the ADC or unconjugated MMAE were identified based on age (24 to 90 years), sex, race (Caucasian, Asian, or Black), renal impairment and mild hepatic impairment (total bilirubin of 1 to 1.5 × ULN and AST any, or total bilirubin ≤ULN and AST >ULN)"*

Another approach, albeit less successful to date but with huge potential for the future, is PBPK modeling. PBPK models are being used to simulate a drug's pharmacokinetics in certain populations that might otherwise be difficult to study clinically. For example, Novartis used PBPK to simulate the effect of varying degrees of renal and hepatic impairment on their drug asciminib (Scemblix). Although the FDA ultimately rejected their proposal that the PBPK model was sufficient to support label claims, this demonstrates that companies are beginning to leverage new technologies to address this difficult-to-study subpopulation.

Although this section has focused on renal impairment, most of the information applies to studies in patients with hepatic impairment. Of course, volunteers with hepatic impairment do not use creatinine clearance to assess disease severity. Instead, a composite score like the Child-Pugh score may be used, a multi-factor score based on the volunteer's bilirubin, albumin, INR, the severity of

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

their ascites, and any encephalopathy they may exhibit. Study designs, analysis, and interpretation follow similar practices as renal impairment.

Information on the effects of renal or hepatic impairment is presented in Section 8 of the label. An example of how such information is conveyed is in the Mounjaro label:

> *8.6 Renal Impairment*
>
> *No dosage adjustment of MOUNJARO is recommended for patients with renal impairment. In subjects with renal impairment including end-stage renal disease (ESRD), no change in tirzepatide pharmacokinetics (PK) was observed [see Clinical Pharmacology (12.3)]. Monitor renal function when initiating or escalating doses of MOUNJARO in patients with renal impairment reporting severe adverse gastrointestinal reactions [see Warnings and Precautions (5.5)].*
>
> *8.7 Hepatic Impairment*
>
> *No dosage adjustment of MOUNJARO is recommended for patients with hepatic impairment. In a clinical pharmacology study in subjects with varying degrees of hepatic impairment, no change in tirzepatide PK was observed [see Clinical Pharmacology (12.3)].*

In this instance, no dose adjustment is needed in patients with any degree of renal or hepatic impairment.

The Elderly

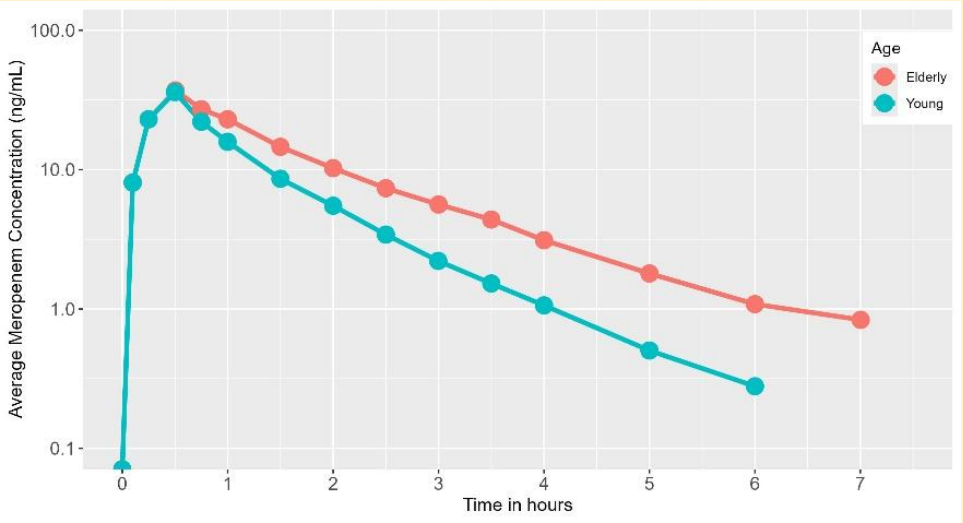
The elderly are underrepresented in clinical trials because few enroll. This is problematic because the global population is aging and living longer. The World Health Organization estimates that 1 in 6 people will be 65 or older by 2030. It has also been estimated that the elderly use more than one-third of all prescribed medications [273]. Despite these statistics, many elderly patients are excluded from clinical trials, either by direct age exclusions or indirectly through exclusion criteria that are affected by age, like having a participant's renal function be above some arbitrary cutoff (renal function declines with age; that's normal, not a disease state) or a limit on the number or use of concomitant medications. Often, elderly patients are excluded because of perceived vulnerability in this population that may skew the efficacy and safety of a clinical trial.

We all know that our bodies change as we age, so it should come as no surprise that the pharmacokinetics and pharmacodynamics of drugs also change. These changes can impact all pharmacokinetic processes, including absorption, distribution, metabolism, and excretion, as well as the pharmacodynamics of the drug at the receptor level. In general, with aging, there is a:

- decrease and slowing of drug absorption,
- changes in the percent fat and body water, which may alter a drug's volume of distribution,
- a decrease in serum albumin concentration, which may also alter a drug's volume of distribution, mainly for acidic drugs,
- decreased Phase 1 metabolism with little change in Phase 2 metabolism,
- reduced hepatic blood flow and a decrease in liver mass, which may decrease the liver clearance of the drug,
- reduced renal function, which may decrease the renal excretion of drugs,
- generally increased half-life for many drugs,
- with most elderly tending to take many medications (polypharmacy), the risk of drug interactions is considerably higher than in younger patients, and
- increased sensitivity to medications.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Example of Age-Related Changes in Pharmacokinetics



Eight healthy subjects were administered a single intravenous dose of meropenem, an antibiotic. Elderly volunteers, aged 67 to 80 years, had higher concentrations during the terminal elimination phase and a longer half-life than young volunteers, aged 30 to 34 years, due to reduced renal and hepatic clearance in the elderly group. Note that the plot is displayed on a log scale to emphasize the longer half-life in elderly volunteers. Figure redrawn from [274].

There are also changes at the receptor level after accounting for drug concentrations. For instance, the elderly have increased sensitivity to benzodiazepines like diazepam [276], which affect GABA receptors, whereas they have reduced sensitivity towards isoproterenol and propranolol at beta-receptors [277]. All in all, there is a general tendency for the elderly to require lower doses than their younger counterparts.

Other things that should be considered when prescribing to the elderly are that drug therapy tends to start at lower doses than in younger patients, frequent dose adjustments may be needed initially, there is an increased risk of drug interactions, compliance is an issue as patients may forget to take their medication or may not take their medication as prescribed, and their physical condition may limit what medication they can swallow, e.g., large antibiotic pills may be problematic in the elderly.

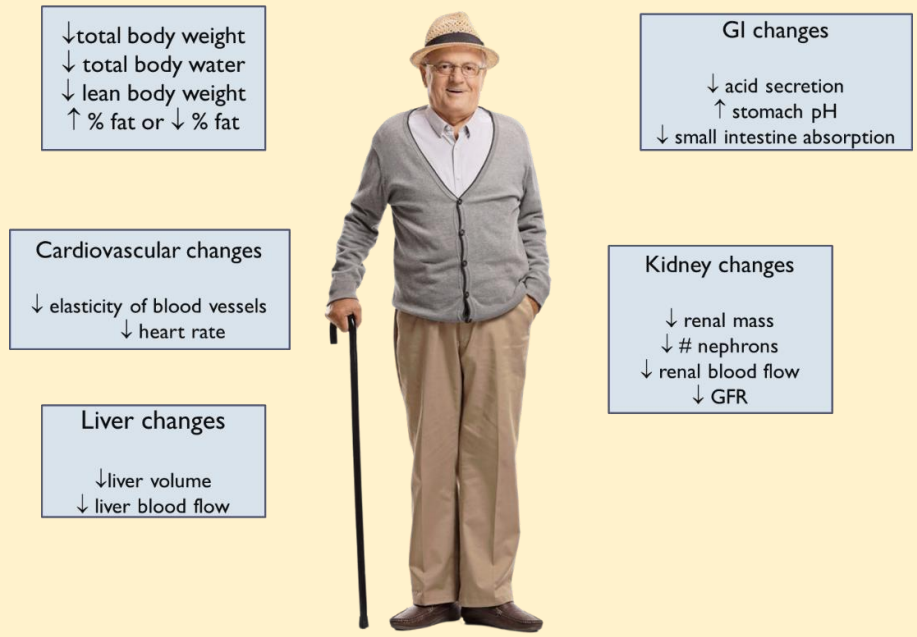
Partial List of Drugs That Show Changes in Sensitivity in the Elderly

Drug	Effect	Change
Diazepam	Sedation	↑
Diltiazem	Blood Pressure	↑
Furosemide	Diuretic	↓
Morphine	Analgesia	↑
Propranolol	Blood pressure	↑
Scopolamine	Cognitive function	↓
Verapamil	Blood pressure	↑
Warfarin	Anticoagulant	↑

Reference: [275].

Although there is a need for better dosing information in the elderly, and the FDA has numerous guidances to encourage inclusion of elderly patients in clinical trials, there are few instances of drug

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



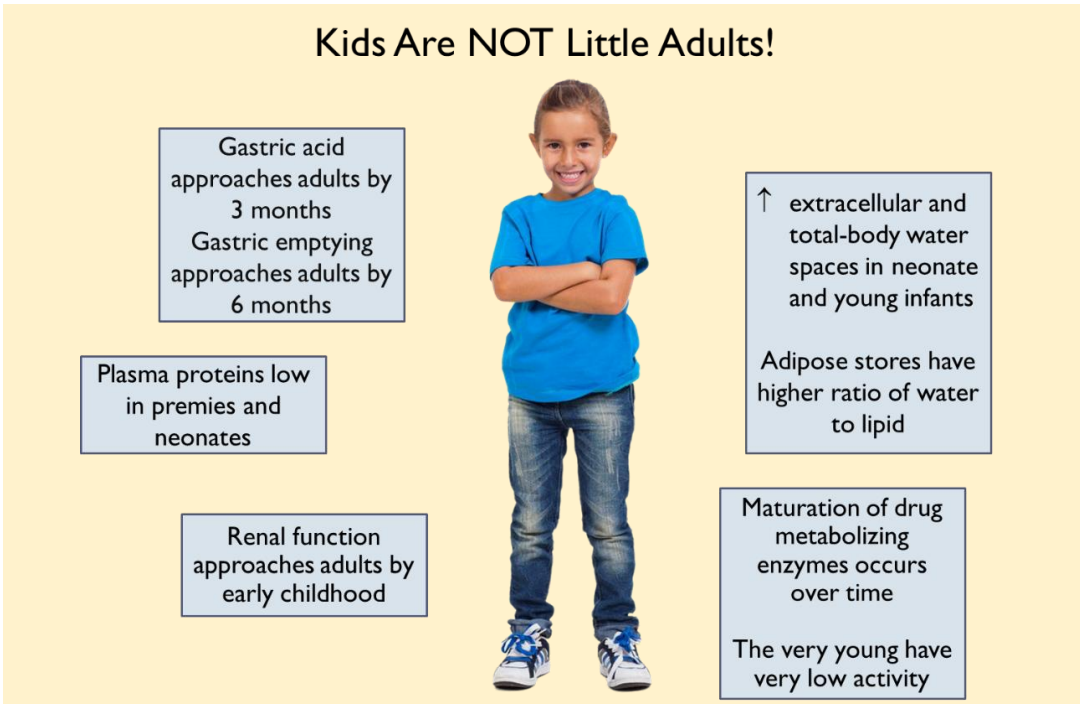
Many physiological changes occur in the elderly, which may affect a drug's pharmacokinetics and pharmacodynamics.

labels that identify the elderly as a special population. An example of labeling in the elderly is the label for lamorexant (Dayvigo), used to treat insomnia in adults. The elderly section says:

> *Of the total number of patients treated with DAYVIGO (n=1418) in controlled Phase 3 studies, 491 patients were 65 years and over, and 87 patients were 75 years and over. Overall, efficacy results for patients < 65 years of age were similar compared to patients ≥65 years.*
>
> *In a pooled analysis of Study 1 (the first 30 days) and Study 2, the incidence of somnolence in patients ≥65 years with DAYVIGO 10 mg was higher (9.8%) compared to 7.7% in patients <65 years. The incidence of somnolence with DAYVIGO 5 mg was similar in patients ≥65 years (4.9%) and < 65 years (5.1%). The incidence of somnolence in patients treated with placebo was 2% or less regardless of age [see Clinical Studies (14.2)]. Because DAYVIGO can increase somnolence and drowsiness, patients, particularly the elderly, are at a higher risk of falls [see Warning and Precautions (5.1)]. Exercise caution when using doses higher than 5 mg in patients ≥65 years old.*

This label does not explicitly state that no dose change is necessary, but it does imply that one is not required. Additional information is provided throughout the label. Section 12.3 on Pharmacokinetics states that age does not affect lamborexant pharmacokinetics. Section 5.1 on CNS Depressant Effects states: "*Because DAYVIGO can cause drowsiness, patients, particularly the elderly, are at a higher risk of falls.*" Also, because this drug affects the central nervous system, a study assessed the effect of lamborexant on driving ability. Section 14.2 states, "*Although DAYVIGO at doses of 5 mg and 10 mg did not cause statistically significant impairment in next-morning driving performance in adult or elderly subjects (compared with placebo), driving ability was impaired in some subjects taking 10 mg DAYVIGO*." Overall, for this particular drug, there does not appear to be a difference in effects

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Children differ from adults in physiology, which may lead to differences in drug pharmacokinetics and pharmacodynamics.

between the elderly and the young, nor in pharmacokinetics.

Pediatrics

Children are not little adults. The younger children are, the more physiologically different they are from adults. Because children are constantly changing as they get older, they tend to be grouped into more homogeneous populations:

- Preterm newborn,
- Newborn (0-28 days),
- Infant (> 28 days to 12 months),
- Toddler (> 12 months to 23 months),
- Preschool child (2 years to 5 years),
- School-age child (5 years to 11 years), and
- Adolescents (12 to 18 years)

Due to these physiological differences, numerous examples exist of drugs whose pharmacokinetics differ between children and adults. As in adults, it is important to find the right dose for the right patient; the same holds for children. Sponsors must find the right dose for children across all the different age groups for whom the drug. The bottom line for children is that smaller individuals require smaller absolute doses; however, when doses are expressed as mass per unit weight, e.g., mg/kg, children tend to require higher doses than adults.

There are many reviews on the physiological differences between adults and children [278, 279], and this topic will not be extensively reviewed here. Suffice it to say that after children are born, it takes time for many physiological functions to mature to adult levels. Pediatric dosing was introduced in the section on Post-Marketing Requirements and Phase 4 Studies. It is generally accepted that dosing

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Bottom Line

Smaller People Need Smaller Absolute Doses

but dose/weight is greater for children than adults

regimens that account solely for differences in dosing based on weight or body surface area are relatively accurate down to around 2 years of age. Older rules, like Clark's rule or Salisbury's rule, are no longer considered acceptable [280]. A common rule still being recommended is BSA-based dosing:

$$\text{Pediatric Dose} = \frac{\text{Child BSA in m}^2}{1.73\ \text{m}^2}\text{Adult Dose}\ . \tag{37}$$

The dose is then rounded to the nearest available strength formulation. This approach is quite intuitive and is a linear extrapolation. If the dose for a normal-sized adult is 100 mg, then the dose for a child half that size is half the adult dose, or 50 mg. One problem with this approach is that it requires an estimate of BSA, which complicates the calculation. BSA-based dosing also overestimates the dose in younger children and underestimates the dose in older children. Further, if only a limited number of dose strengths are available, as is often the case with solid dosage forms like pills or tablets, this complicates matters. For example, what if no 50 mg dose had been available? What if the only available dose was 100 mg? Then what?

Due to the limitations of older rules of thumb, which are only valid down to around 2 years of age, today, when estimating dose requirements in children, more accurate but correspondingly more complex methods are employed to predict dosage requirements. These methods are explained in the chapter on Drug Development, using Cresemba as an example. Simply put, down to 2 years of age, clinical pharmacologists use allometric scaling of pharmacokinetic parameters, like clearance and volume of distribution, to estimate the values of those parameters in children and then simulate a drug's pharmacokinetics in children by varying the dose until drug concentrations similar to those seen in adults are observed. This method assumes that the pharmacodynamics in children are similar to those of adults. If this is not the case, then those differences must also be accounted for in the dose estimation. More specialized modeling is required for children under 2 years old. PBPK models, which account for enzyme maturation, renal and hepatic function, differences in fat and water compartments, and changes in protein binding, are typically used to predict doses in this population.

Frequently, for children, particularly oral doses, a table is given with the corresponding dose for a given child's weight. For example, shown below is the dosing table for acetaminophen in children:

Table 15
Acetaminophen Dosing Recommendations for Children

Child's weight (pounds)	6-11	12-17	18-23	24-35	36-47	48-59	60-71	72-95	96+	lbs
Liquid 160 mg/ 5 milliliters (mL)	1.25	2.5	3.75	5	7.5	10	12.5	15	20	mL
Liquid 160 mg/ 1 teaspoon (tsp)	--	½	¾	1	1½	2	2½	3	4	Tsp
Chewable 160 mg tablets	--	--	--	1	1½	2	2½	3	4	Tabs
Adult 325 mg tablets	--	--	--	--	--	1	1	1½	2	Tabs
Adult 500 mg tablets	--	--	--	--	--	--	--	1	1	Tab
Reference: www.tylenol.com (accessed 29 July 2024)										

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

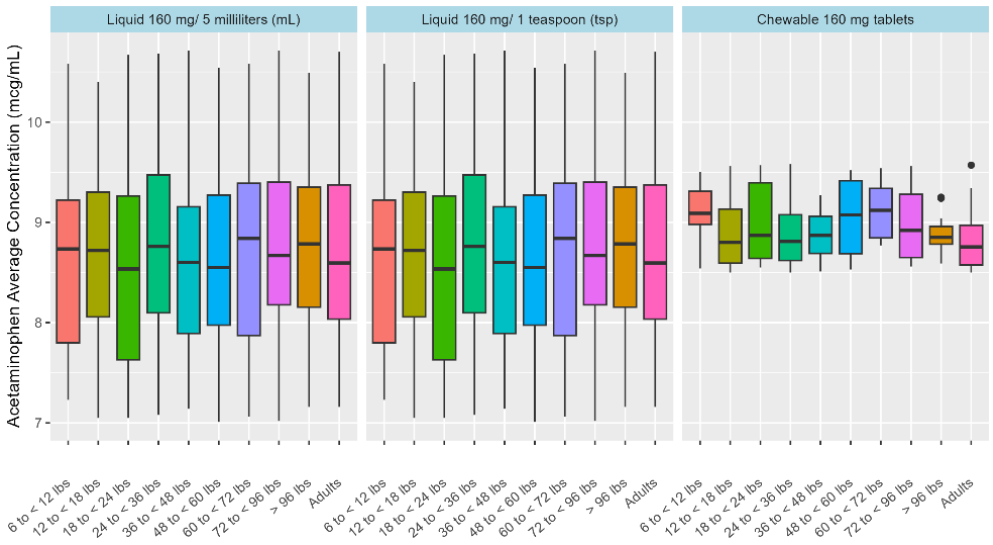


Figure 75: Simulated acetaminophen concentrations in children and adults for different weight groups using the acetaminophen dosing table. Children and adults were simulated across the different weight groups, and acetaminophen concentrations were simulated using the model reported by [281], then summarized using a box-and-whisker plot. Each box contains 75% of the simulated values. The line in the middle of the box is the median. The whiskers contain approximately 95% of the simulated values. The therapeutic range of acetaminophen is between 5 and 20 mcg/mL.

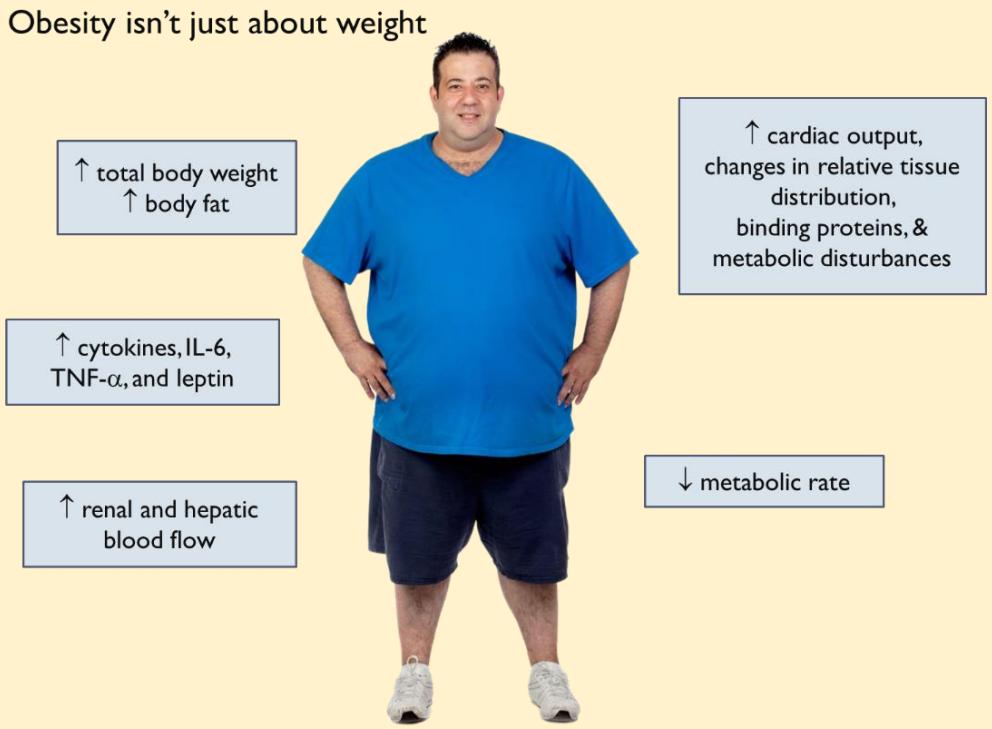
You will notice that tablets are generally unavailable until the child is around 4 years of age because very young children have difficulty swallowing pills. For this reason, a suspension formulation is usually prepared for very young children. In the absence of a suspension, parents may be asked to crush the tablet or capsule and mix it into applesauce or baby food.

But how good is this dosing table? A good dosing recommendation should result in similar drug concentrations across different weights and, for acetaminophen, fall within the therapeutic window of 5 to 20 mcg/mL. A simulation was conducted to test how good the dosing recommendations are in Figure 75. Average concentrations were simulated after a single acetaminophen dose using the pharmacokinetic model reported in [281] with the doses recommended in the table. A total of 100 children were simulated within each weight range used in the table. For simplicity, only the liquid and chewable tablet formulations were examined. The results shown in Figure 75 are quite good. For each formulation, the average concentration across all weight ranges tested was similar to, or not very different from, concentrations observed in adults. Within a formulation, there was little variability across weight groups. Furthermore, the average concentrations were within the therapeutic window of 5 to 20 mcg/mL for all formulations across all weights. Hence, the dosing table is quite good at generating consistent acetaminophen concentrations across a wide range of doses and different formulations.

Obesity

Today, obesity is a pandemic of worldwide proportions. Almost a billion people worldwide are considered obese. The Centers for Disease Control estimates that 4 in 10 adult Americans were obese in 2020 [282], while the World Health Organization estimates worldwide this number is closer to 1 in 8 [283]. Obesity is not just an increase in weight; it increases the risk of other diseases like high

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Obesity changes metabolic processes and anatomic alterations, affecting a drug's pharmacokinetics and pharmacodynamics.

blood pressure, diabetes, and cancer, as well as psychosocial illnesses like depression. Adult obesity is a body mass index (BMI) of greater than 30 kg/m^2, while overweight is a BMI of 25 to 30 kg/m^2. BMI is not the same as body weight. The two are correlated, but not identical. BMI attempts to adjust for the fact that taller people tend to be heavier. But it is not perfect – far from that. BMI has been criticized as a measure of obesity for many reasons, such as identifying athletes like football and rugby players as being obese, i.e., it treats people with high muscle mass as obese. Nevertheless, it is still the measure almost exclusively used to exclude patients from clinical trials.

Obesity is not just a condition of increased body fat. Obesity is a complex disease that is multifactorial in its etiology and results in metabolic and anatomic alterations [284]. Obesity results in endocrine changes that increase cytokine and other hormone levels, leading to a metabolic state that mimics a low-grade inflammatory state. As body fat increases, the cardiovascular system is burdened, leading to increased cardiac output and subsequent changes in renal and hepatic blood flow. Over time, this increased burden on the cardiovascular system increases the risk of heart failure and stroke. Long-term obesity can lead to non-alcoholic fatty liver syndrome (NAFLD), which can further lead to changes in drug metabolism. Also, hypertension can occur over time, which can affect kidney function, further leading to changes in drug excretion.

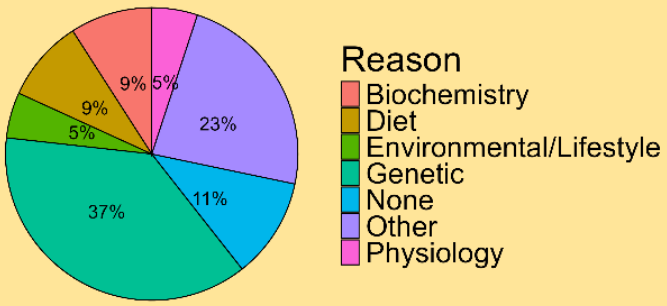
The physiologic alterations that occur in obesity can affect many pharmacokinetic processes, the most obvious of which may be the volume of distribution due to increased fat deposits. However, alterations in the drug's volume of distribution depend on its physicochemical properties. Lipophilic drugs tend to have a larger volume of distribution in obese patients because of increased drug distribution into fat, but lipophobic (hydrophilic) drugs have no such alteration; lipophilic drugs remain in the blood. Biologics, such as monoclonal antibodies, have a higher volume of distribution because they are confined to the blood compartment. With obesity, although there is an increase in blood volume, blood volume does not change to the same extent as total body weight [285].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Ethnicity Question

Special populations arise through exclusion criteria or underrepresentation in clinical trials. Could ethnicity, like being Japanese, lead to a special population status? Yes, first-generation Japanese patients may be underrepresented in clinical trials and may require dose adjustments compared to their Western counterparts. In one study comparing warfarin pharmacokinetics in Japanese and Western patients, profound differences in pharmacokinetics were observed, leading to large pharmacodynamic differences:

Parameter	Caucasian Patients (n=47)	Japanese Patients (n=90)
Unbound plasma concentration (ng/mL)	4.75 ± 2.49	2.05 ± 0.92
Body-weight normalized apparent oral CL (mL/min/kg)	5.10 ± 3.10	11.2 ± 5.43
Body-weight normalized metabolic clearance (mL/min/kg)	0.010 ± 0.009	0.030 ± 0.040
INR (a measure of the drug's effect; an INR of 2 to 3 is considered acceptable for an anticoagulant)	2.82 ± 1.21	1.58 ± 0.43



Other factors, such as genetic differences or dietary variations, may explain the variations in pharmacokinetics between Westerners and individuals from other ethnicities. This pie chart represents one research group's attribution of percentage variance to factors that might explain inter-ethnic differences [286].

Although there were differences in CYP2C9 genotype frequency between Japanese and Westerners, this could only partially explain the differences [287]. Although the authors do not make recommendations regarding dose modifications in Japanese, it is interesting to note that a retrospective analysis of drug labels in Japan showed that for 1 in 3 drug labels, the maximum approved dose in the United States was 2-times higher than the maximum approved dose in Japan and that 1 in 3 drug labels had a different approved dose than in the United States, mostly often being smaller than in the United States. This example is representative of what is seen when other ethnicities are examined. Other ethnicities, like Chinese, could also have been used, as many studies have found pharmacokinetic and pharmacodynamic differences across a wide range of ethnicities [286].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Increased cytokine levels can lead to downregulation of cytochrome P450 3A4 and reduced drug metabolism. This effect has been noted in obese patients [288], although other enzymes, such as cytochrome P450s CYP1A2 and CYP2D6, may show increased activity with obesity. As with blood volume, the kidneys and liver increase in size with obesity, but these increases are not proportional to total body weight. Hence, there might be increases in drug clearance with obesity, but the change may not reflect the magnitude of the increase in total body weight. In one study, even though obese patients had almost twice the average weight as healthy normal volunteers, their glomerular filtration rate, one of the major processes the kidney uses to eliminate drugs and metabolic waste from the body, was only increased by 62% [289].

Even though clearance and volume of distribution tend to be higher in obesity, it is difficult to generalize statements like 'obesity results in higher dose requirements' or 'obesity requires lower dose requirements' because the drug itself plays a role in this regard. For example, anesthetics used in surgery tend to be very lipophilic because they need to distribute into the brain quickly for their sedative effects to occur rapidly. Anesthetics, like propofol, are frequently dosed per body weight, e.g., mg/kg, and, as such, obese patients are at risk of overdose. Hence, anesthetics are difficult to dose for several reasons. They show prolonged accumulation in fat relative to the brain, which makes it difficult to achieve sedation, and they show slower recovery because, as the drug redistributes from fat to blood, it distributes to the brain, prolonging recovery [291]. As such, obese patients tend to have lower anesthetic dose requirements. On the other hand, aminoglycoside antibiotics, like tobramycin or gentamycin, are almost exclusively cleared by the kidneys. Obese patients require higher aminoglycoside dose requirements than patients with normal weight, even after controlling for other factors like age and sex [292]. For example, obese patients require a gentamycin dose of around 540 mg/day to have similar drug concentrations as normal-weight patients administered a dose of 380 mg/day, presumably because of the increase in glomerular filtration rate seen in obese patients.

Pharmacokinetic changes observed in obese patients can result in pharmaco-dynamic alterations. Holt et al. [290] reported that the effectiveness of oral contraceptives is decreased in obese patients. The risk of pregnancy in patients taking oral contraceptives was 60% higher in patients having a BMI > 27.3 kg/m^2 and 70% higher in patients with a BMI > 32.2 kg/m^2. Furthermore, emergency contraceptives, like NorLevo, began losing effectiveness at 75 kg (165 lbs.) and showed an absence of effectiveness at 80 kg (176 lbs.).

Cancer drugs are another class of drugs that are often dosed per body weight, either mass-scaled to total body weight or mass-scaled to body surface area (BSA), i.e., mg/kg or mg/BSA. Strictly following a weight-based approach without consideration of the properties of the drug may result in overdosing. As such, physicians may reduce the dose out of concern for their patient's safety, but in so doing, may underdose the patient and reduce the drug's efficacy [293]. One analysis estimated that around 40% of obese cancer patients are underdosed. To correct for total body weight, physicians may use an alternate weight measure like ideal body weight or lean body mass, which are theoretical weights given a patient's sex and height. There is no guarantee that these alternative weight measures produce better outcomes in obese patients. Even with some of the newer oral tyrosine kinase inhibitors, which are flat-based dosed, i.e., doses are not based on weight and are fixed in nature, e.g., 300 mg once-daily, evidence suggests these drugs may also need dose adjustments because of obesity [294]. However, current guidelines by the American Society of Clinical Oncology recommend no dose adjustment for obese patients for any drug and to dose per the recommended dose for a normal weight patient[295]. This guideline was based on a literature review of more than 60 research articles that showed that obese patients tolerate cancer drugs as well as normal-weight patients, regardless of the higher administered dose.

Some drugs have what is called a "dose cap," where once a dose is reached, no further increases in

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

dose can occur. For example, enfortumab vedotin (Padcev) is dosed at 1.25 mg/kg (up to a maximum of 125 mg) as an intravenous infusion over 30 minutes on Days 1 and 8 of a 21-day cycle, until disease progression or unacceptable toxicity is observed. The dosing cap is 125 mg; no one can receive more than 125 mg, even if they weigh more than 100 kg. This dose cap was chosen based on the number of serious drug-related adverse events seen in the early trials with the drug. In early trials, the incidence of Grade 3 or higher hyperglycemia was greater in obese patients compared to normal-weight patients, even after controlling for prior history of hyperglycemia or pre-existing diabetes at baseline. Four of 5 deaths in an early clinical trial with enfortumab vedotin, before the dosing cap was instituted, occurred in obese patients. Based on these adverse events, a dose cap was implemented, and no further incidences of drug-related deaths were observed in later trials [296].

Despite what is known about obesity, few drug labels have information about dosing modifications in obese patients. In an analysis of 185 pediatric drugs approved between 2016 and 2021, only 13% had information related to obesity, mostly referring to adverse events or some underlying comorbidity, in their label. Only 7% had information on pharmacokinetics in obese patients [297]. Only one drug (Saxenda, liraglutide) had any dose modification for obese patients. An example of how obesity is described in the pharmacokinetics section of the drug label is Cubicin (daptomycin):

> *The pharmacokinetics of daptomycin were evaluated in 6 moderately obese (Body Mass Index [BMI] 25 to 39.9 kg/m^2) and 6 extremely obese (BMI ≥40 kg/m^2) adult subjects and controls matched for age, gender, and renal function. Following administration of CUBICIN by IV infusion over a 30-minute period as a single 4 mg/kg dose based on total body weight, the total plasma clearance of daptomycin normalized to total body weight was approximately 15% lower in moderately obese subjects and 23% lower in extremely obese subjects than in nonobese controls. The $AUC_{0-\infty}$ of daptomycin was approximately 30% higher in moderately obese subjects and 31% higher in extremely obese subjects than in nonobese controls. The differences were most likely due to differences in the renal clearance of daptomycin. No adjustment of CUBICIN dosage is warranted in obese patients.*

Lovenox (enoxaparin) has an interesting label. Not only is there a section for obese patients in Section 8, but there is also a section for low-weight patients:

> *8.8 Low-Weight Patients*
>
> *An increase in exposure of enoxaparin sodium with prophylactic dosages (non-weight adjusted) has been observed in low-weight women (< 45 kg) and low-weight men (< 57 kg). Observe low-weight patients frequently for signs and symptoms of bleeding [see Clinical Pharmacology (12.3)].*
>
> *8.9 Obese Patients*
>
> *Obese patients are at higher risk for thromboembolism. The safety and efficacy of prophylactic doses of Lovenox in obese patients (BMI >30 kg/m^2) has not been fully determined and there is no consensus for dose adjustment. Observe these patients carefully for signs and symptoms of thromboembolism*

It is ironic that 'Obesity' is a special population when more than 40% of Americans are considered 'obese.' Therefore, the general lack of information in most drug labels on how to dose obese patients is somewhat surprising. From personal experience, every drug I have worked on that has been approved has had a physician inquire with the company after approval as to how to dose the drug in obese patients.

Summary

Pharmaceutical marketing groups would love it if one dose applied to everyone – one dose to rule them all (in Lord of the Rings jargon). But that often does not apply. Certain groups of individuals may require dose modifications, such as patients with altered physiology due to disease or other

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

comorbidities. Or not. Sometimes, physicians want to know if they can use the label dose in their patient, maybe someone who has mild hepatic impairment, because they have hepatitis. Furthermore, underrepresented patients, whether excluded directly or indirectly due to enrollment criteria or poor participation in clinical trials, may require special attention from physicians. For these reasons, these patients are put in the Specific Populations section of the label to help physicians make informed dosing decisions for their patients.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Chapter 9

Drug Interactions

A drug interaction occurs when some factor (called the perpetrator[7]) increases or decreases the concentration or effect of a drug (called the victim). The factor causing the interaction can be smoking, food, a beverage, or another drug, for example. The factor could also be the drug you are developing; it may cause a drug interaction with another drug. The major cause of drug interactions is an interaction between two drugs, leading to what are called drug-drug interactions (DDIs). The incidence of DDIs is unknown, but estimates suggest that the prevalence rate for hospital admissions due to a DDI ranges from approximately 1% to about 3.5%, with about 10% of patients experiencing a DDI while hospitalized [240, 299]. If you consider that about 30 million people are hospitalized per year in the United States, then 300,000 to 1 million people are hospitalized for having a DDI. The exact number of DDIs that occur outside the hospital each year is not known, but to say that they are in the millions is probably not a far reach. These are staggering numbers, so it's understandable that drug companies do a lot of work to understand where, with what, and under what conditions these interactions may occur. This chapter will discuss drug interactions by presenting case studies and examining how drug companies assess the potential for interactions with new drugs being developed today.

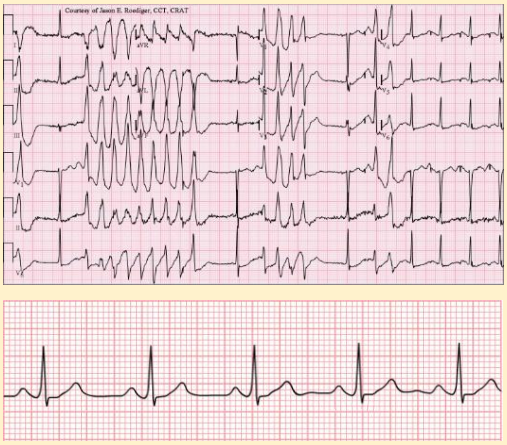
Example of a Drug-Drug Interaction: The Terfenadine Story

Seldane (terfenadine) was approved in 1985 as the first non-sedating antihistamine for the treatment of allergies. Ironically, it was originally developed as a tranquilizer but was found to be inactive because it did not reach the brain. After examining its chemical structure and recognizing the structural similarity to diphenhydramine, an antihistamine, terfenadine was also found to possess antihistamine properties. After approval, Seldane would quickly go on to become one of the earliest drugs to reach a billion dollars in sales in a year and may have been responsible for direct-to-consumer advertising (but that is a different story).

By the end of the 1980s, several case reports were published on possible drug-drug interactions with terfenadine. Furthermore, MedWatch, the FDA's monitoring system for detecting adverse drug reactions, began to gather evidence of adverse events in patients. Some of the patients were taking other drugs like erythromycin, while some had evidence of liver impairment, both of which might inhibit terfenadine metabolism. Some of these cases were associated with sudden death with cardiac

[7] Recent regulatory guidances are changing the terms 'perpetrator' and 'victim' to 'precipitant' and 'object', respectively, perhaps in a misguided attempt to move past those original pejorative terms and for political correctness 298. Food and Drug Administration, *Guidance for Industry: Drug Interaction Studies*. 2024 (https://www.fda.gov/media/161199/download)..

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026



Example of Torsades de Pointes on an ECG (top) and normal ECG (bottom). Torsades de pointes is associated with prolonged QT interval, characterized by twisting QRS complexes, potentially causing syncope or sudden cardiac death Figures reprinted from Wikimedia Creative Commons.

involvement. All evidence pointed towards this being an issue that somehow affected the heart.

One such case report was published by Monahan et al. [241] of a 39-year-old female with a 2-day history of intermittent syncope and lightheadedness. Ten days earlier, she had started taking Seldane 60 mg twice daily and Ceclor (cefaclor) 250 mg three times daily for a recurrent sinus infection. On day 8, she developed a urinary tract infection (UTI). She stopped taking Ceclor and self-treated with leftover ketoconazole 200 mg twice daily from a previous infection. Shortly after that, she had a couple of episodes of heart palpitations and syncope, and when presented to the emergency room, her ECG showed Torsades de Pointes (French for twisting of the points), a type of cardiac arrhythmia where the lower chambers of the heart do not beat properly. If left untreated, Torsades can be fatal.

We have all seen an ECG measuring the heart's electrical activity (see Figure 63). The different ups and downs are associated with distinct electrical events in the heart, which are designated by names such as the PR or QT interval. The PR interval corresponds to the depolarization of atrial cells in the heart. In contrast, the QT interval is associated with the depolarization and repolarization of the ventricles in the heart. The time sum of all the intervals represents one heartbeat. Because the total interval length can vary with heart rate, some intervals are corrected to control for it. QT is one of these intervals. In a normal person, the QTc interval (the 'c' indicates the interval was corrected for heart rate) ranges from about 360 to 460 milliseconds. In the patient in question, her QTc interval was 655 msec. Physicians discontinued all medications, and within 7 days, her QTc interval decreased back to a normal value of 429 msec. She was discharged without further incident.

Based on reports like these, it became apparent that terfenadine was affecting the heart's ion channels. But why? Terfenadine is itself a prodrug; it has no activity on its own. It is, however, completely and rapidly metabolized by CYP3A4 to fexofenadine, which accounts for all the drug's activity. While fexofenadine is metabolized by a small amount by CYP3A4, its disposition is largely driven by the transporters organic anion transporting polypeptide (OATP) and P-gp, both of which are inhibited by terfenadine. A normal person should have little to no detectable terfenadine in the blood. In the case study above, plasma concentrations of terfenadine and fexofenadine were 57 ng/mL and 385 ng/mL, respectively, both of which were significantly higher than normal. It turns out that drugs like ketoconazole, which the patient was taking, and erythromycin, as well as medical conditions that result in liver impairment, will inhibit the metabolism of drugs by CYP3A4 and inhibit transporters, like OATP and P-gp. Taken together, greater-than-normal terfenadine and fexofenadine concentrations are observed.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Before Allegra, Before Claritin, there was Seldane, the First Non-Sedating Antihistamine

Seldane was also the first drug removed from the market because it could cause sudden cardiac death in some patients, which was later found to be due to blockade of potassium ion channels (HERG channels) in the heart. Other drugs were removed for the same reason, including cisapride, astemizole, and grepafloxacin

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

But why did these higher-than-normal concentrations prolong the QT interval and lead to syncope, and even possibly death, in some patients? During a heartbeat, ions flow across the cell membrane through ion channels, resulting in the peaks and troughs seen on an ECG. Sodium and calcium flow into the cell, while potassium flows out. One of these ion channels, through which potassium flows, is the HERG channel, also known as the human ether-a-go-go channel. Terfenadine blocks this channel, prolonging the QT interval. If the QT interval is sufficiently prolonged, it may affect the next heartbeat, and if this happens repeatedly, the heart may go into arrhythmia, leading to syncope and possibly death.

Side note: After Seldane was removed from the market, the FDA evaluated all approved drugs for others that might cause QT prolongation. Two were identified: cisapride and astemizole. Both were subsequently removed from the market.

The FDA issued a black box on the Seldane label in 1992, warning physicians about the possible rare cardiovascular adverse events that may occur with Seldane use. The manufacturer, Hoechst Marion Roussel (HMR), also sent a "Dear Doctor" letter to every physician in the United States, once in 1990 and again in 1992, warning them of the risks associated with combining Seldane with antifungals or certain antibiotics, and with administering it to patients with liver impairment. By 1996, however, the manufacturer and FDA agreed the risk outweighed the benefit (stopping a runny nose vs. possible death) and that Seldane should be removed from the market, which it was in 1998.

Example of a Drug-Beverage Interaction: Fexofenadine and Grapefruit Juice

The story of Seldane doesn't end in 1998. The manufacturer of Seldane, Hoechst Marion Roussel (HMR), recognized the implications of the black box warning and that Seldane is a prodrug, so why not develop the metabolite as a replacement product? And that is indeed what happened. The metabolite, fexofenadine, was soon developed. With Seldane, HMR knew which fexofenadine concentrations were necessary for activity and quickly identified the doses required to achieve those concentrations in patients. From there, it was a quick path to approval. Perhaps the one hiccup in the development of fexofenadine was proving that there was no cardiac liability like terfenadine had. HMR conducted one of the first studies, now known as a "thorough QT study," and proved that fexofenadine did not affect cardiac conduction. Fexofenadine was approved in 1996 and marketed as Allegra, still used today as an over-the-counter medication.

But Allegra is not without its own drug interactions. One of those is with fruit juices. You shouldn't take Allegra with grapefruit or orange juice. This interaction seemed so odd at the time that the sponsor took the unusual step of airing a television commercial to notify the public.

How some fruit juices interact with drugs was discovered purely by accident. In 1989, pharmacologist David Bailey and colleagues [300] were studying the effect of alcohol on the antihypertensive effects of felodipine. Patients were randomized to felodipine alone or felodipine coadministered with alcohol, and then after a washout period, the opposite treatment was administered. Each patient would receive both treatments. To blind the patients as to whether they were receiving felodipine alone or felodipine with alcohol, the investigators added grapefruit juice to mask the distinctive taste of alcohol. They found that felodipine concentrations were higher than expected in both groups (Figure 76), an effect they attributed to grapefruit juice.

But they couldn't be sure that the effect was caused by grapefruit juice; it could have been due to something else. They also wondered if grapefruit juice affected other drugs. They then conducted a better-controlled experiment, ensuring the effect was due to grapefruit juice. Again, using a crossover design in which each patient received one of the treatments in each period, Bailey and colleagues studied felodipine alone with water, felodipine with grapefruit juice, and felodipine with orange juice.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

They repeated the results and demonstrated conclusively that grapefruit juice was the culprit. Felodipine AUC and C_{max} were 2- to 3-fold higher when administered with grapefruit juice than water. Orange juice had no effect. They then studied whether grapefruit juice would affect nifedipine, a felodipine analog. Nifedipine AUC and C_{max} were increased by 25% to 50% when administered with grapefruit juice, not as much as felodipine concentrations, but increased nevertheless [301].

Partial List of Drugs Whose Absorption is Affected by Grapefruit Juice

- Maraviroc
- Statins like lovastatin and atorvastin
- Ergotamine
- Oral Ketamine
- Lurasidone
- Many cancer drugs like crizotinib, dasatinib, and erlotinib
- Everolimus
- Tacrolimus
- Erythromycin
- Amiodarone
- Oxycodone
- Dextromethorphan

What could be causing this increase? A later study by a different research group showed that the increase in exposure from grapefruit juice did not occur following intravenous administration [302]. Thus, the increase in exposure was due to an alteration in the amount of drug absorbed after oral administration. We now know that grapefruit juice affects many enzymes and transporters. Felodipine is a substrate of both p-glycoprotein and CYP3A4, both of which are inhibited by grapefruit juice. Normally, most of the drug is not absorbed and bioavailability is low. With grapefruit juice, oral bioavailability is increased.

Grapefruit juice has since been shown to affect the absorption of many drugs, sometimes leading to severe adverse reactions. Drugs with already high bioavailability (> 70%) have a relatively small interaction with grapefruit juice, whereas drugs with low bioavailability (< 10%) are the most affected. Furthermore, repeatedly drinking grapefruit juice can increase the effect compared to a single glass, suggesting a cumulative effect. But none of this explains 'how.' How does grapefruit juice affect these proteins? In the case of CYP3A, the components in grapefruit juice do not simply compete with the drug for CYP3A-mediated metabolism. With grapefruit juice, the effect is more insidious.

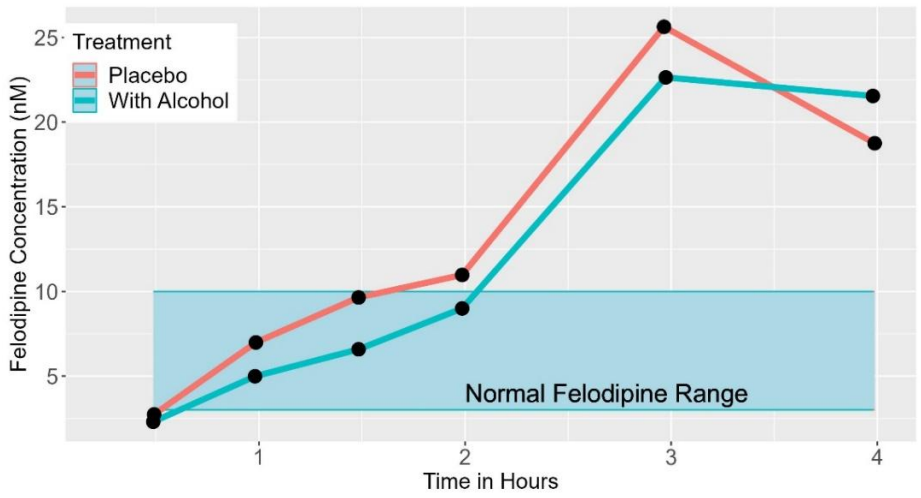


Figure 76: Felodipine concentrations in mildly hypertensive patients alone and in combination with alcohol (top). Both groups received grapefruit juice as a masking agent. Felodipine concentrations were much higher than expected in both groups. Figure redrawn from [300].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

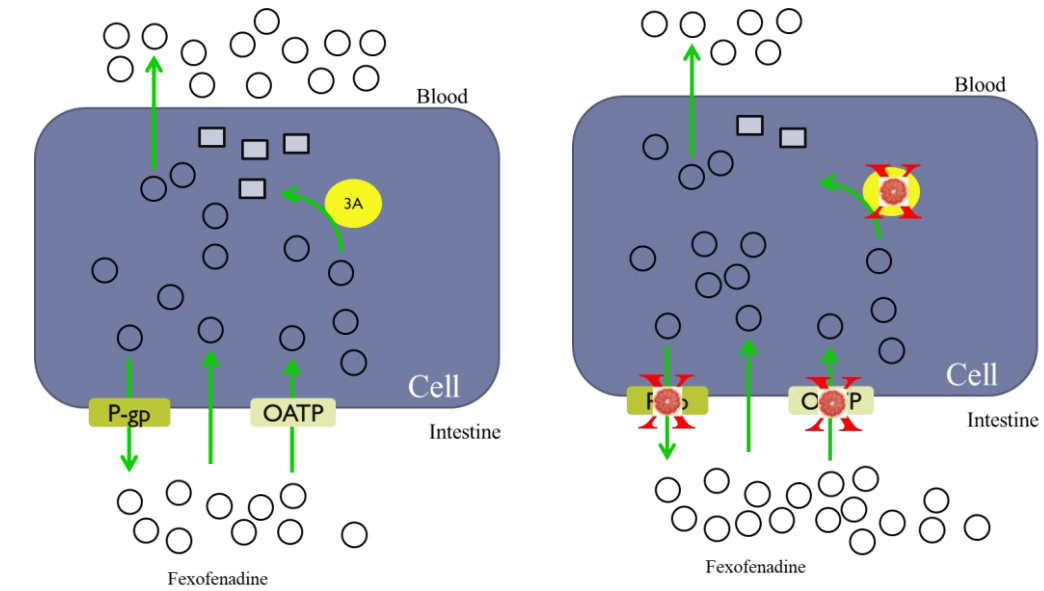


Figure 77: Events within the gastrointestinal tract alone and in the presence of grapefruit juice. Under normal conditions, fexofenadine is absorbed into the cell, either by diffusion or via a transporter such as the organic anion transporter polypeptide (OATP), where it is metabolized by CYP3A or transported back into the gut lumen by p-glycoprotein (P-gp). Any drug not metabolized in the cell or transported out of the cell is available for absorption into the blood (left figure). Grapefruit juice inhibits P-gp, OATP, and CYP3A, but primarily OATP. As a result, less drug is absorbed into the body and plasma fexofenadine concentrations are lower (right figure).

The components of grapefruit juice irreversibly inhibit CYP3A; the enzyme only recovers after new enzyme is synthesized. Studies have shown it takes about a week for CYP3A to recover from a single glass of grapefruit juice [303]. Surprisingly, liver CYP3A is unaffected by grapefruit juice, probably because the components in grapefruit juice that cause the interaction do not get absorbed and therefore does not make it to the liver.

Similar effects are seen with the transporters P-gp and OATP. Not only do the components of grapefruit juice block these transports, but there is also evidence that grapefruit juice decreases the number of transporters on the cell membrane [304] (Figure 77). The exact mechanism for grapefruit juice's effect on these transporters was unknown for a long time. But in 2007, the primary chemical component in grapefruit juice that causes the effect was identified as naringin, the chemical also responsible for its bitter taste.

Fexofenadine is a good substrate of the influx transporter OATP, which increases the amount of fexofenadine absorbed by the body. When healthy volunteers were administered fexofenadine and a relatively low dose of naringin simultaneously, fexofenadine concentrations were lower compared to simultaneous coadministration of fexofenadine and water [305]. Fexofenadine concentrations were even lower when grapefruit juice was coadministered than when naringin alone was administered. This is why you should not consume fruit juices with Allegra, as stated in the commercial: your fexofenadine concentrations will be much lower than expected, and the drug may not work as intended.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Table 16 Naringin Concentration (in mg/dL) for Different Brands of Grapefruit Juice	
Kroger (unsweetened grapefruit juice from concentrate)	63.8
Ocean Spray (grapefruit juice from 100% concentrate)	51.2
Tropicana (pure premium red ruby grapefruit – the original, not the concentrate)	51.2
Texsun (100% pure pink grapefruit juice from concentrate)	46.3
Kroger (unsweetened grapefruit from concentrate)	38.5
Kroger (grapefruit from concentrate unsweetened)	35.3
Ocean Spray (pink grapefruit concentrate)	19.1
Kroger (ruby red grapefruit juice cocktail from concentrate contains 35% juice)	16.3
Ocean Spray (ruby red grapefruit concentrate from cocktail)	14.6
Data reprinted from [306].	

After the discovery of grapefruit juice's effect, drug companies started to do drug interaction studies with grapefruit juice, but this was abandoned after a few years because the studies were shown to be difficult to reproduce and the results were erratic. This was because the different brands of grapefruit juice contained different concentrations of naringin (Table 16). Even lots from the same manufacturer contained different concentrations [307]. Instead, drug companies switched to using a drug that had the same effect but whose dose and purity could be controlled. Originally, ketoconazole was used as a strong inhibitor of CYP3A and P-gp in DDI studies in humans. However, due to ketoconazole's potential hepatotoxicity, itraconazole has been used more recently because it has a better safety profile.

After Bailey's discovery, others started to look for other drug-juice interactions. There have been case reports of carbamazepine, an antiepileptic medication, bioavailability and toxicity being increased after drinking lime juice and of a patient dying from fatal hemorrhage after taking warfarin and cranberry juice together for 2 weeks. Based on clinical studies, it has been demonstrated that:

- Plain orange juice will decrease atenolol, alendronate, and the fluoroquinolone antibiotics;
- Seville orange juice can increase felodipine exposure;
- Pomelo juice can increase cyclosporine exposure but will decrease sildenafil exposure;
- Grape juice can decrease cyclosporine exposure;
- And there are many others [308].

Many fruits and fruit juices can have a pharmacokinetic interaction with drugs

The bottom line with fruit juices is that all of them have the potential to affect a drug's pharmacokinetics, and patients should check with their doctor or pharmacist for potential interactions.

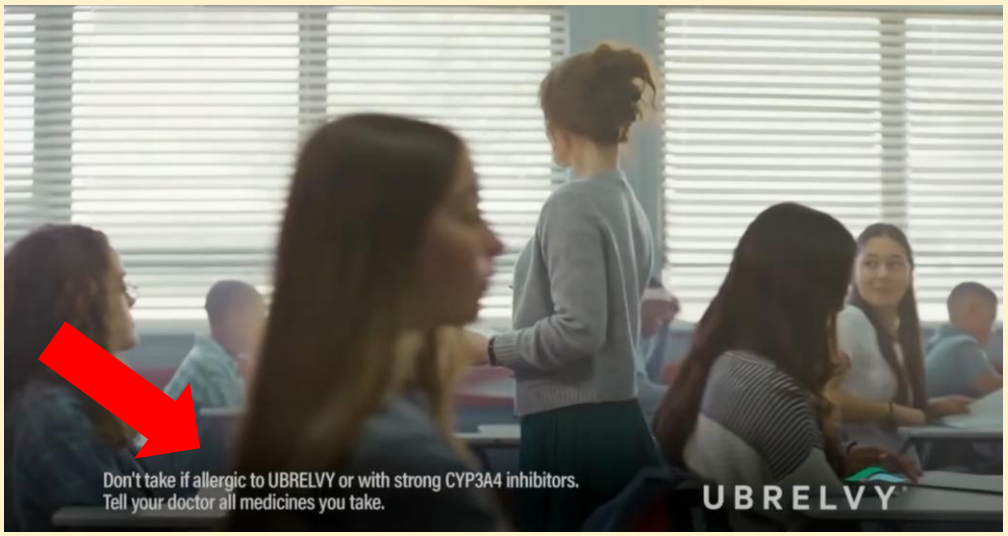
Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Figure 78: St. John's wort increases P-gp expression and decreases plasma concentrations for many drugs. The top graphic shows the effect of St. John's wort in the small intestine. The net effect of p-glycoprotein induction is decreased bioavailability and reduced drug concentrations (bottom). The bottom figure shows mean alprazolam concentrations alone and after 14 days of St. John's wort administration in 12 healthy volunteers. Figure redrawn from [309].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

What is a Strong CYP3A Inhibitor?

I was watching television, and a drug commercial for Ubrelvy (Ubrogepant) came on. Ubrelvy is used to treat acute episodes of migraine. In the middle of the commercial, the announcer says, 'Don't take … with strong CYP3A inhibitors,' and I thought to myself, "Who is going to know what that means"? Television commercials are aimed at the public, and this was a highly technical phrase. How many people will know what a strong CYP3A inhibitor is?



UBRELVY® (ubrogepant) tablets 50mg 100mg

The enzyme CYP3A4 almost exclusively metabolizes Ubrogepant to an inactive metabolite. Drugs that are metabolized primarily by a single pathway are candidates for drug interactions with metabolic inhibitors of that pathway. Inhibitors of CYP3A are classified as strong, moderate, or weak by how much they increase drug concentrations. Typically, strong inhibitors increase drug concentrations by more than 5-fold, moderate inhibitors by 2- to 5-fold, and weak inhibitors by < 2-fold. For Ubrogepant, ketoconazole, a prototypical strong CYP3A inhibitor, increased AUC by almost 10-fold and Cmax by more than 5-fold. Verapamil, a prototypical moderate inhibitor of CYP3A, increased AUC by 3.5-fold and Cmax by 2.8-fold. By definition, strong inhibitors have the largest drug interactions and are a cause for concern. Ubrogepant is available in 50 and 100-mg tablets. The label for Ubrelvy states that strong CYP3A inhibitors should be avoided. For patients taking a moderate CYP3A inhibitor, use the 50 mg dose and avoid a second dose within 24 hours. Patients taking weak inhibitors should use the 50 mg dose, but may take another dose within 2 hours.

But what drugs are CYP inhibitors? Indiana University maintains a list of known inhibitors and inducers for various metabolic enzymes, called the Flockhart Table, in honor of David Flockhart, the man who started and maintained it for many years. Typical strong inhibitors include antifungal drugs such as itraconazole and ketoconazole, carbamazepine, eliglustat, omeprazole, triazolam, and trazodone. Moderate CYP3A inhibitors include diazepam, hydrocortisone, ondansetron, and pantoprazole. Now, when you see these commercials, you will know what they mean.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Example of a Drug-Supplement Interaction: St. John's Wort

Supplements were discussed in the chapter What is a Drug? St. John's wort is a plant-derived supplement promoted for treating various ailments. The first case report of possible problems with St. John's wort was from a 29-year-old female who received a cadaveric kidney transplant and was taking cyclosporine to prevent rejection [310]. The patient had stable organ function and had stable cyclosporine blood concentrations.

The patient began to self-medicate with St. John's wort, and after 4 to 8 weeks, her cyclosporine concentrations began to decline and became subtherapeutic. The patient stopped St. John's wort and, by 4 weeks, reattained therapeutic concentrations, but by then, the patient rejected the kidney and had to return to dialysis. A second case report in a 55-year-old kidney transplant patient soon followed. These reports suggested that St. John's wort may be interacting with cyclosporine. We now know that St. John's wort interacts with many, many drugs, including antidepressants, anticancer drugs, digoxin, warfarin, simvastatin, and oral contraceptives.

The hypericum perforatum plant

St. John's wort is derived from the flowers, stems, and leaves of the Hypericum perforatum plant. It can be turned into a pill, tea, or oil and is used for a variety of indications, including gastrointestinal illnesses, menopause, and depression. It derives its name from St. John the Baptist because it blooms around the time of the festival in his name, in late June.

P-glycoprotein has been discussed many times. It is a transporter that spans the cell membrane and effluxes drugs from the inside of the cell to the outside. It is an evolutionary detoxification mechanism that prevents the absorption of toxins and other harmful substances. Drugs that are P-glycoprotein substrates and get inside the cell are actively removed from the cell by the transporter. Repeated and prolonged administration of St. John's wort causes the body to make more p-glycoprotein in the cell membrane, a process known as induction. St. John's wort increases the amount of drug removed from the cell, thereby reducing the drug's bioavailability and decreasing blood concentrations as a result (Figure 78). St. John's wort also increases CYP3A expression, creating a kind of perfect storm. With St. John's wort, less of the drug is absorbed, and the drug that is absorbed has greater metabolism, both of which decrease drug concentrations. Everyone agrees – taking St. John's wort is a bad idea if you take any other medication because it is likely to reduce the efficacy of the other medications you are taking.

Example of a Drug-Smoking Interaction: Tarceva

Tarceva (erlotinib) was approved in 2004 for the treatment of non-small cell lung cancer patients. In the Phase 3 pivotal trial with erlotinib, patients with a history of smoking were less likely to respond to treatment than non-smokers. Further analysis showed that trough plasma concentrations in smokers were half those of non-smokers. CYP3A4, CYP1A2, and CYP2C8 metabolize Erlotinib. Smoking is known to increase the metabolism of many drugs by its ability to induce their metabolism [311]. Smoking has a particularly strong effect on CYP1A2. To confirm that there was a pharmacokinetic interaction with smoking, a Phase 1 study was conducted in 12 healthy male volunteers who were non-smokers for at least a year or had a history of smoking 10 or more cigarettes per day for at least 4 months [312]. Smokers had concentrations that were significantly lower than those of non-smokers at both doses (Figure 79). Hence, smoking's effect on erlotinib is

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

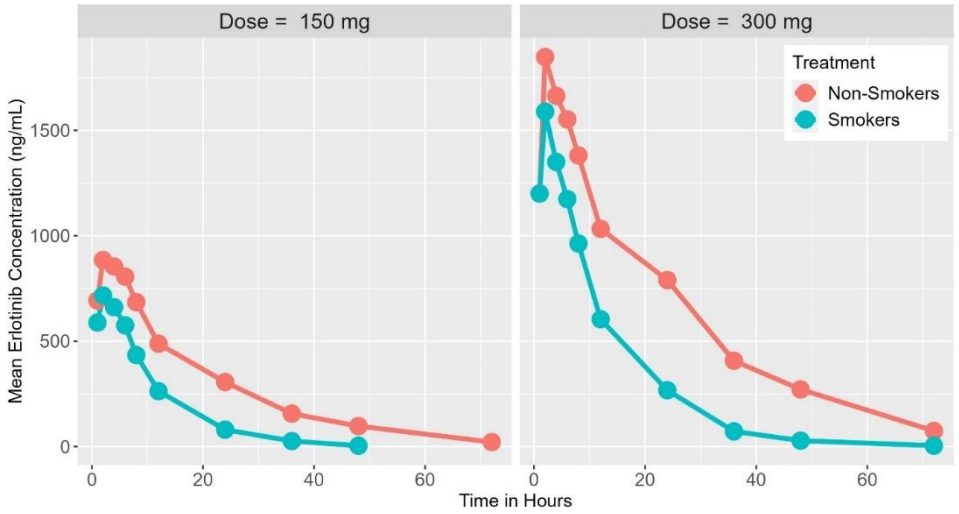


Figure 79: Erlotinib concentrations are decreased in smokers compared to non-smokers. The plot shows mean plasma erlotinib concentrations in 12 healthy volunteers who were smokers or nonsmokers. All subjects received a single 150 mg dose of erlotinib on Day 1, followed by a single 300 mg dose on Day 15. Plasma samples were collected for 72 hours and analyzed for erlotinib concentration. Figure redrawn from [312].

consistent with its effect on CYP1A2 induction. Smoking can affect other drugs as well, such as caffeine, diazepam, insulin, heparin, and theophylline, to name a few.

Example of a Drug-Food Interaction: Tykerb

We have all had experiences where our pharmacist specifically told us to take a particular medication with food or to make sure we take it on an empty stomach. Food intake increases, decreases, or has no effect on the absorption of orally administered drugs. Because of this, it is particularly important to understand the effect of food on a new drug's pharmacokinetics.

For example, consider Tykerb (lapatinib), which is used in combination with capecitabine or letrozole to treat breast cancer. The label for Tykerb states that:

> *"TYKERB should be taken once daily," and "TYKERB should be taken at least one hour before or one hour after a meal."*

Why is it necessary to take Tykerb on an empty stomach? In a Phase 1 study, 27 patients received either lapatinib 1500 mg after an overnight fast, immediately after a low-fat breakfast, or immediately after a high-fat breakfast on three separate occasions, one week apart [313]. The low-fat meal consisted of Special K cereal, skim milk, toast with jam, apple juice, and a cup of coffee or tea (520 calories). The high-fat breakfast consisted of eggs, bacon, hash browns, whole milk, and coffee or tea (1036 calories). Figure 80 shows the results of the study. Both low-fat and high-fat breakfasts significantly increased lapatinib concentrations. AUC and C_{max} increased 3- to 4-fold compared with the fasting state, potentially increasing the risk of adverse events.

Surprisingly, the food effect associated with lapatinib proved quite controversial. At the time, lapatinib was an expensive drug—a month's supply could cost about $3000. The controversy was started by Ratain and Cohen [314] when they wrote an opinion piece in the Journal of Clinical Oncology, the leading oncology journal worldwide, read by thousands of oncologists every month, stating that, contrary to what was on the label, it would be to the benefit of patients if they took their

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

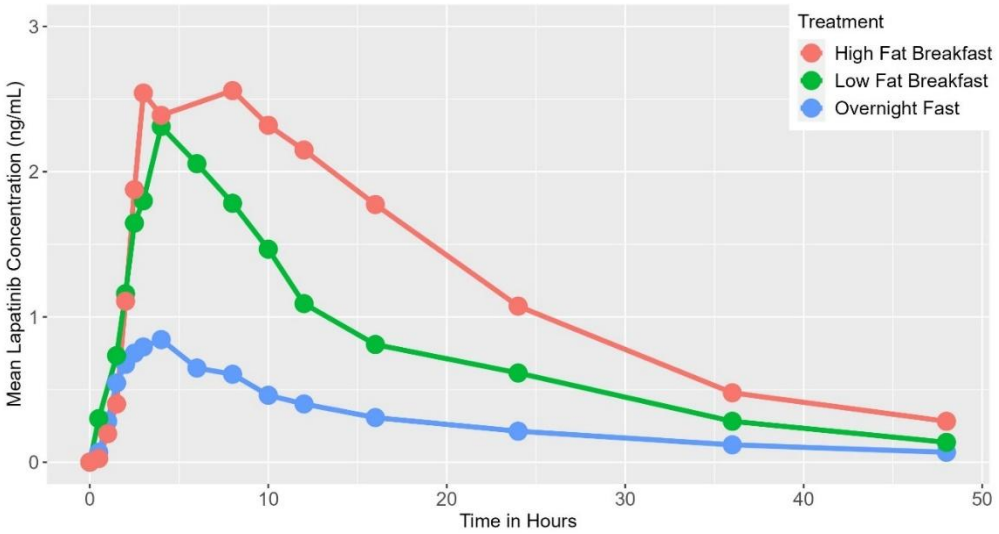


Figure 80: Food increases lapatinib bioavailability. The plot shows mean lapatinib concentrations in 27 cancer patients administered 1500 mg lapatinib after an overnight fast, immediately after a low-fat meal, or immediately after a high-fat meal. Figure redrawn from [313].

lapatinib with food. They could potentially save $1,740 per month, or 60%. In their Phase 3 studies, GlaxoSmithKline instructed patients to take lapatinib after a 4-hour fast. The label reflects the findings from the pivotal clinical studies that led to approval. However, two prominent oncologists argue that the label is wrong, to the detriment of patients' wallets. This was followed by rebuttals from the company and other oncologists [315, 316], stating Ratain and Cohen were wrong. Yes, on average, concentrations are higher, but this comes at the expense of increased variability. The variability in lapatinib concentrations among patients taking the same dose was large. GSK believed the best way to minimize variability was to administer the drug while the patient was fasting. Taking it with food would result in much wider variability, which would produce erratic concentrations from day to day, potentially affecting efficacy and/or safety. Ratain and Cohen [317] held fast to the criticism of GSK and chastised the FDA for their labeling of lapatinib and the increased risk of QT prolongation with lapatinib when administered with food. Ultimately, the label remained unchanged, and the controversy eventually subsided.

How food affects the pharmacokinetics of a drug is complicated [318]. Food slows gastric emptying from the stomach to the small intestine, raises the pH to near-neutral levels, increases blood flow to the liver, and slows transit through the small and large intestines. The magnitude of these changes depends on the meal's composition, whether it is largely carbohydrates, fats, or a mix, whether it is mostly solids, soup-like, or a liquid, and the size of the meal. The net effect is that the amount of drug absorbed, the rate of drug absorption, or both can be affected.

Scientists have gotten reasonably good at predicting, at least qualitatively, how food will affect the absorption of a drug [118] using the two primary events necessary for oral drug absorption to occur:

1. Dissolution of the drug from its formulation into solution within the gastrointestinal tract (solubility); and
2. Absorption of the drug across the gastrointestinal tract cells from the lumen inside the gut to the blood (permeability).

These two parameters form the basis of the Biopharmaceutics Classification System, which was originally proposed to indicate whether a bioequivalence study is necessary for a drug having an immediate-release solid oral dosage form [319]. If you classify drugs as high and low across these

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

criteria, drug absorption is generally affected as follows:

Table 17
Effect of Food on Drug Absorption

	Solubility	
Permeability	**High**	**Low**
High	Class I AUC unchanged	Class II ↑ AUC
Low	Class III ↓ AUC	Class IV ? AUC

Class I drugs exhibit maximal absorption at a fast rate. Using this simple table, the effect of food on a drug's AUC can be accurately predicted 60% to 70% of the time. Using more complicated statistical methods can improve this value to about 80% [320].

Example of a Drug-Protein Interaction: Tocilizumab

An interesting example of a drug interaction is between tocilizumab and drugs metabolized by cytochrome P450. When monoclonal antibodies were first developed in the early 2000s, it was thought that their potential for interactions with other drugs would be low because of their high specificity for their targets. And for the vast majority of monoclonal antibodies, this was true. For many cytochrome P450 enzymes, inflammation and infection downregulate their expression, reducing drug metabolism. Cytokines are chemicals secreted by immune cells that affect other cells. Examples included tumor necrosis factor (TNF) and interleukins (IL). In patients with autoimmune disorders, such as rheumatoid arthritis (RA), numerous cytokines are elevated in plasma, including interleukin-6 (IL-6), which has multiple effects on the body, including stimulating T-cell activation and promoting blood cell growth and differentiation. IL-6 has also been shown to reduce the activity of cytochrome enzymes in vitro, albeit at concentrations much higher than those reported in RA patients, which raises the potential for IL-6 to affect the metabolism of many drugs.

Simvastatin is a prodrug that is used to treat high cholesterol. It is hydrolyzed through CYP3A4 to its active β-hydroxy simvastatin acid metabolite, which is also metabolized through CYP3A4. Hence, simvastatin is an excellent probe for CYP3A4 activity. Reduced CYP3A4 activity is expected to increase both simvastatin concentrations and acid metabolite concentrations. Tocilizumab (Tofidence) is an IL-6 monoclonal antibody against the IL-6 receptor and is used to treat adult and juvenile RA. Even though IL-6 concentrations are increased in RA, tocilizumab blocks IL-6's effect by inhibiting the IL-6 receptor, and reduces inflammation. Tocilizumab also blocks the in vitro downregulation of CYP enzymes in human hepatocytes in the presence of IL-6, suggesting that tocilizumab may block the downregulation of CYP enzymes in humans and lead to:

↑ tocilizumab → ↓ IL-6 activity → ↑ CYP3A → ↓ drug concentrations.

In a clinical study, 12 RA patients received simvastatin on Day 1, 15, and 43 with a single infusion of tocilizumab administered on Day 8. Simvastatin concentrations were measured on Day 1, 15, and 43 (Figure 81). Observed concentrations for both parent and metabolite on Day 15 were lower than on Day 1 and remained lower on Day 43, 35 days after administration of tocilizumab [321]. This change in drug concentration was mirrored by a decrease in plasma C-reactive protein, a biomarker for inflammation. Similar results were seen with omeprazole, a marker of CYP2C19 metabolism.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

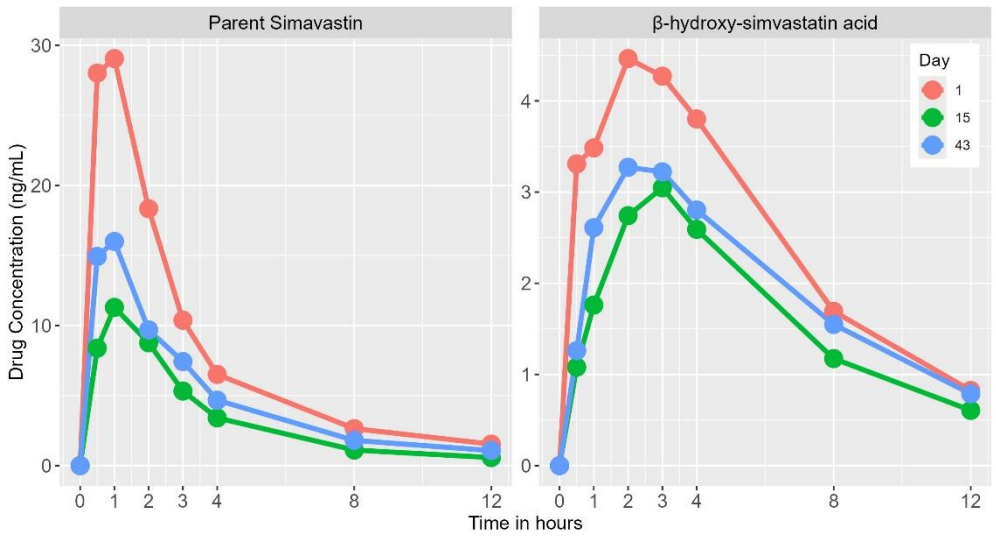
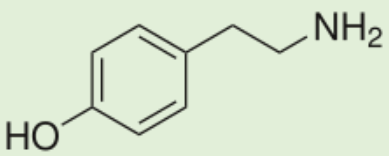


Figure 81: Simvastatin and β-acid metabolite concentrations on Days 1, 15, and 43 in patients with rheumatoid arthritis. A single dose of tocilizumab was administered on Day 8. Figure redrawn from [321].

These results suggest that tocilizumab may indirectly affect the metabolism of CYP-mediated drugs, particularly those with a narrow therapeutic window. The package insert for Tofidence states that physicians *"exercise caution when coadministering Tofidence with CYP3A4 substrate drugs where a decrease in effectiveness is undesirable, e.g., oral contraceptives."* We now know this drug interaction can occur with any drug that modulates cytokine concentrations. For instance, peginterferon alfa-2b, a proinflammatory cytokine used in patients with hepatitis, doubled dextromethorphan (a probe for CYP2D6) concentrations in healthy subjects but not in patients with hepatitis. Additionally, monoclonal antibodies, such as blinatumomab and mosunetuzumab, can transiently increase cytokine levels, potentially reducing the risk of drug interactions when cytokine modulation is involved.

Example of a Drug-Food Interaction: MAO Inhibitors

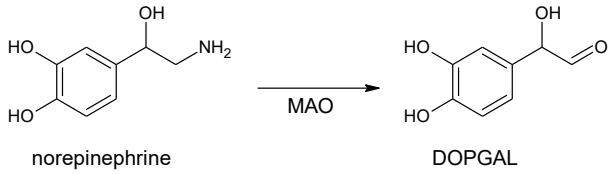
Not all interactions are pharmacokinetic; some DDIs affect pharmacodynamics without affecting a drug's pharmacokinetics. One example is the interaction between MAO inhibitors and certain foods. In the 1950s, it was believed that depression was due to a deficiency of the neurotransmitters serotonin, norepinephrine, and dopamine in the brain. This was referred to as the monoamine theory of depression. Scientists believed that if they could boost the levels of these neurotransmitters in the synapses of the neurons in the brain, then a patient's depression could be cured. Monoamine oxidase (MAO) is the enzyme responsible for metabolizing many neurotransmitters, including monoamines thought to be involved in depression.



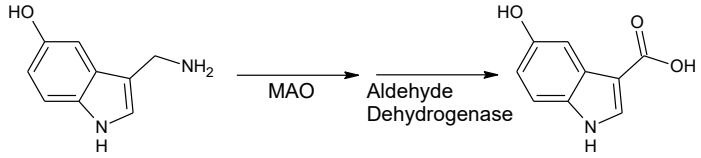
Structure of tyramine, a neurotransmitter that increases blood pressure, is notably similar to other neurotransmitters like norepinephrine, dopamine, and serotonin.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

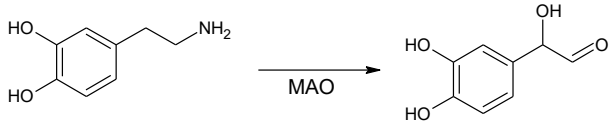
The metabolism of norepinephrine, serotonin, and dopamine is shown below:



norepinephrine DOPGAL



Serotonin (5HT) 5-HIAA



Dopamine DOPAC

MAO inhibitors were discovered by accident. Iproniazid was originally developed to treat tuberculosis. It was found that in patients with tuberculosis and depression, their depression was alleviated after treatment with iproniazid. It was soon found that iproniazid irreversibly inhibits MAO. Inhibiting neurotransmitter metabolism increases neurotransmitter levels in synaptic clefts, thereby boosting their effect. This initiated the search for other MAO inhibitors. During the late 1950s to 1970s, several MAO inhibitors were introduced, including moclobemide, tranylcypromine, selegiline, and phenelzine, which were more selective and did not irreversibly inhibit MAO.

The utility of these drugs was limited by their side effects and the potential for drug dependence-like withdrawal syndrome when administration is stopped. One of these side effects is a hypertensive crisis that can happen due to certain types of food interactions. MAO not only metabolizes norepinephrine, serotonin, and dopamine, but also metabolizes other amines, such as phenylethylamine, benzylamine, epinephrine, and tyramine. Tyramine acts to release neurotransmitters from their vesicles inside neurons to the synapse. It is found in certain types of food, such as aged cheeses (Swiss, Parmesan, Cheddar, etc.), alcohol (like beer and wine), cured meats (like salami), smoked meats (like hot dogs), soy sauce, soybeans, or overripe fruit. Tyramine increases the amount of neurotransmitters in the synaptic space, and MAO inhibitors then inhibit the metabolism of those neurotransmitters, enhancing their effect, causing a hypertensive crisis. Although the incidence of MAO inhibitor toxicity has decreased significantly in the past 25

Some tyramine-containing foods, like pizza, beer, or salami, interact with MAO inhibitors by inhibiting tyramine metabolism, which may lead to a hypertensive crisis.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

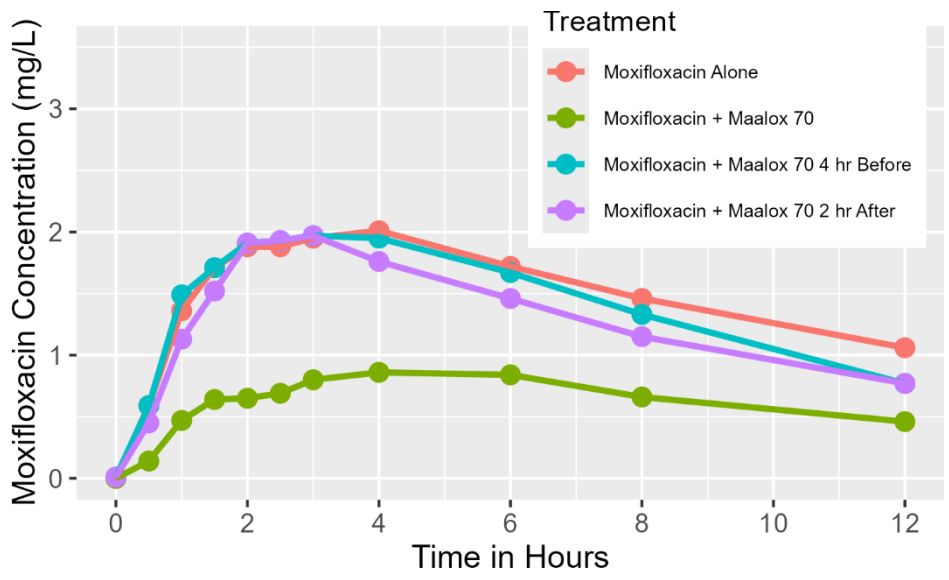


Figure 82: Example of a drug-chemical interaction. The figure shows moxifloxacin concentrations in healthy volunteers alone and with coadministration of Maalox. Coadministration of Maalox results in decreased and slower moxifloxacin absorption. Figure redrawn from [323].

years, there is still about a 1% incidence yearly.

By way of illustration, in normal subjects, tyramine increases systolic blood pressure (SBP) by approximately 30 mmHg in over half the subjects tested [322]. Normal SBP is about 120 mm Hg. However, in patients on MAO inhibitors, only 8 mg of tyramine is needed for a similar increase in SBP. For reference, a slice of cheddar cheese has about 10 mg of tyramine. Hence, in some patients, the simple act of eating a piece of cheese could be sufficient to significantly increase blood pressure to unsafe levels.

Example of a Drug-Chemical Interaction: Antacids

Many people take Tums or Maalox when they have an upset stomach. Neither of these is a drug per se – they are chemicals. Tums is aluminum hydroxide, and Maalox is magnesium hydroxide. When these products enter the stomach, they dissociate into aluminum or magnesium ions and hydroxide ions. The hydroxide ions then interact with the hydrogen ions in the stomach, forming water. This raises the pH of the stomach, stops the acidic feeling, and reduces heartburn. Their mechanism of action is by a chemical reaction. pH ranges from 0 to 14, with '0' being strongly acidic, like battery acid, and '14' being strongly basic, like a drain cleaner; 7 is considered neutral. Water has a pH slightly less than 7. Normal stomach pH is around 1-1.5. Because pH is logarithmic, an increase in pH from 1 to 3 means acid ions have been reduced 100-fold. Antacids can reduce stomach pH from 1.5 to near 7.

Antacids have the potential to affect the oral absorption of drugs [324]. Changing stomach pH can affect the drug's solubility in stomach fluid. Magnesium or aluminum can chelate, or bind, drugs, thereby preventing their absorption. Some antacids may alter gastric motility, potentially slowing or accelerating absorption. Drug companies frequently conduct drug interaction studies with antacids to determine whether an interaction may occur. Not only may they examine the change in pharmacokinetics when a drug is coadministered with an antacid, but they may also look for a window in which it is safe to administer the antacid, say, 2 hours before or after the drug.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

An example of a chemical-drug interaction is shown in Figure 82. Healthy volunteers were administered moxifloxacin alone and in combination with Maalox 70. Maalox was administered either concurrently with moxifloxacin, 4 hours before, or 2 hours after moxifloxacin administration. When moxifloxacin was coadministered with Maalox, moxifloxacin concentrations were reduced by ~60%. If this had been a patient, their moxifloxacin concentration may have been too low to maintain the therapeutic effect of the drug. But when Maalox was administered 4 hours before or 2 hours after, no change in moxifloxacin concentrations was observed. This example underscores the importance of timing and the potential influence of antacids on the pharmacokinetics of an orally administered drug.

Example of a Drug-Pharmacodynamic Interaction: Tirzepatide

GLP-1 agonists are on their way to becoming the best-selling drugs of all time, and Mounjaro (tirzepatide) is the leader at the moment. One effect of GLP-1 agonists is to delay gastric emptying, making patients feel satiated for a longer period, thereby reducing the number of meals and the amount of food eaten. With this pharmacodynamic effect, there is a risk of slowing the absorption of certain drugs; indeed, this is the case with tirzepatide, which has been shown to affect the pharmacokinetics of oral contraceptives. In a study with healthy, female volunteers, the AUC_{ss} and $C_{max,ss}$ decreased by 21% and 55% for ethinyl estradiol and 22% and 55% for norelgestromin, respectively (Table 18).

Table 18
Effect of Tirzepatide on Oral Contraceptive Pharmacokinetics

Parameter	Treatment	Ethinyl Estradiol		Norelgestromin	
		Value	Change	Value	% Change
AUC_{ss} (pg*h/mL)	OC + TZ	811	↓21%	16008	↓22%
	OC alone	1029		20659	
$C_{max,ss}$ (pg/mL)	OC + TZ	50	↓59%	983	↓55%
	OC alone	121		2181	

Taken together, the FDA issued the following recommendations on the label. Within the highlights section is the statement:

> *MOUNJARO delays gastric emptying and has the potential to impact the absorption of concomitantly administered oral medications.*

Within the Use in Specific Population section of the Prescribing Highlights is the statement:

> *Females of Reproductive Potential: Advise females using oral contraceptives to switch to a non-oral contraceptive method or add a barrier method of contraception for 4 weeks after initiation and for 4 weeks after each dose escalation.*

Statements like these are not found in labels for other GLP-1 agonists, like Wegovy (semaglutide), because their effect on oral contraceptives does not appear to be as strong as seen with tirzepatide [325]. But this is not to say that the other GLP-1 agonists don't have drug interactions. They also appear to have them for the same reason. For example, semaglutide affects levothyroxine pharmacokinetics, increasing its AUC by 33% and decreasing its Cmax by 12% [326]. This class of

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

drugs is still so new, it's hard to know all the interactions that may occur with them – time will tell.

Example of a Drug-Microbiome Interaction? Checkpoint Inhibitors and Proton Pump Inhibitors

This interaction is hot off the press and may not even be real, but if it is, its mechanism is very unusual. Checkpoint inhibitors (CPIs) are a type of immunotherapy used to treat cancer. These inhibitors block cell-surface proteins, essentially removing the brakes on the immune system, allowing immune cells to recognize and attack cancer cells more efficiently. There is emerging evidence that proton pump inhibitors (PPIs), such as omeprazole and pantoprazole, worsen survival rates with CPIs. In a meta-analysis of seven studies reporting this finding, use of PPIs was associated with a 12% higher risk of cancer progression and an 18% higher risk of mortality [327].

It has been known for some time that the gut microbiome modulates the efficacy of CPIs, and that PPIs can alter it. Hence, the current theory is that PPIs modulate CPIs through their effect on the microbiome. Whether this is true remains to be seen. Further, all evidence to date has been retrospective analyses; no prospective, randomized clinical trials have been conducted to date confirming this interaction, so it is unclear whether this interaction is real.

Example of a Drug-"Weed" Interaction

Interactions with herbs and plants have already been discussed regarding St. John's Wort. This section, though, will literally be discussing drug interactions with "weed", aka marijuana, and with cannabidiol (CBD). With the rise in marijuana use in the US, as more and more states make it legal to possess (within limits), it is prudent to discuss the possible drug interactions with marijuana and CBD, not only as perpetrators of drug interactions, but also as victims of drug interactions.

Marijuana is the common name for *cannabis*, an herbaceous flowering plant, of which *cannabis sativa* and *cannabis indica* are the most common strains used today. It's unclear exactly how cannabis came to be referred to as marijuana, but today, laws use either cannabis or marijuana interchangeably. Some laws refer to the whole plant, others just to the flower. Some laws refer to marijuana as one part of the cannabis plant, the other part being hemp. Needless to say, it's confusing and complicated, yet weirdly simple because everyone knows what marijuana is.

Cannabis has been used for its various mental and physical effects, as well as for spiritual use, for hundreds of years, and today is used by millions of people worldwide. Although (-)-*trans*-Δ^9-tetrahydrocannabinol, which is often just simplified to tetrahydrocannabinol or THC, is the main active ingredient within cannabis, more than 500 different compounds have been isolated upon analysis after administration into the body, with the major ones being cannabinoids and terpenes. The former provides marijuana's pharmacologic effects, while the latter are responsible for its smell and flavor. Although some cannabinoids, like THC, are psychoactive, most are not, like CBD. With cannabis legalization increasing in the US, manufacturers have been quite creative with regard to route of administration, although inhalational and oral are the major preferred routes. The current legal system for cannabis is complicated because, despite its legalization in many states, cannabis is still a Schedule I drug according to the Federal government. CBD is also complicated; its legality depends on its source, how much THC it contains, and what state and locality you are referring to. CBD is not illegal in small quantities under the 2018 Farm Bill, which legalized CBD and hemp derivatives as long as the product contains no more than 0.3% THC by weight. However, the FDA still prohibits the sale and use of CBD as a dietary supplement or in foods or beverages.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

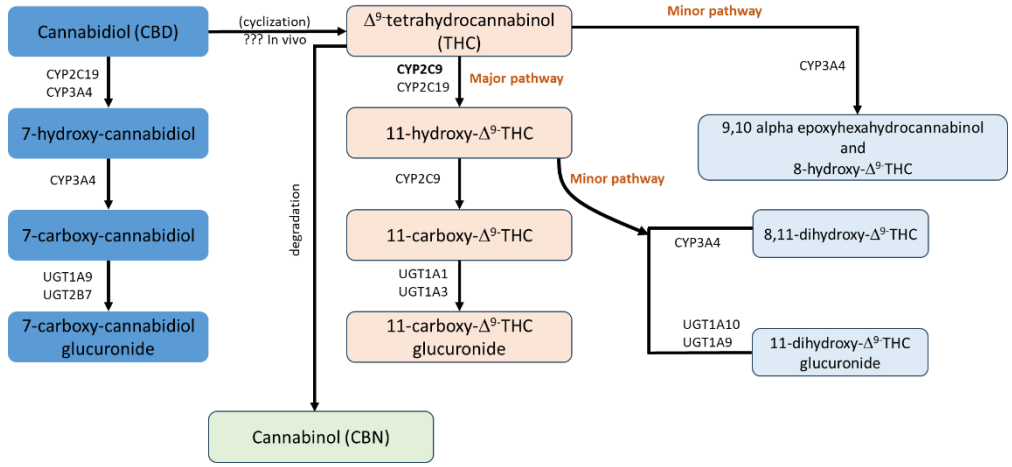
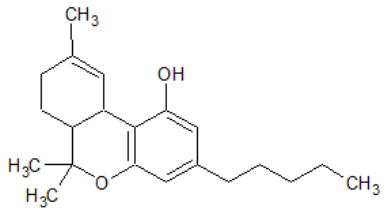
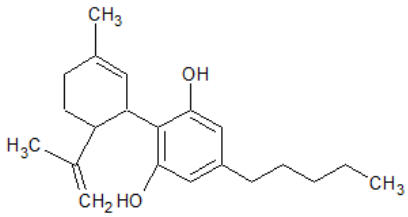


Figure 83: Metabolism of THC and CBD. The conversion of CBD into THC in the body is controversial and there is scientific debate as to whether this occurs. Figure modified from [328].

Structurally, THC and CBD look similar; they have the same number of carbons, oxygens, and hydrogens ($C_{21}H_{30}O_2$), but pharmacologically and metabolically, they are quite different:



Δ9-Tetrahydrocannabidiol (THC)



Cannabidiol

Pharmacologically, THC interacts with endocannabinoid receptors, particularly the CB1-receptor, which is a G-protein receptor found in the brain and peripheral organ, that can modulate mood, memory, appetite, and pain. The metabolism of THC is complicated (Figure 85) and involves many Phase I CYP enzymes, as well as Phase II UGT enzymes [328]. THC has been shown to competitively inhibit several CYP enzymes, including CYP1A2, CYP2B6, CYP2C19, CYP2C9, and CYP2D6, with the greatest inhibition of CYP2C19. THC has also been shown to increase CYP1A2 expression in the liver (induction).

In contrast, the mechanism by which CBD works is unclear. CBD does not interact with endocannabinoid receptors, and yet, CBD has been approved for treatment by the FDA under the brand-name Epidiolex to treat rare forms of epilepsy. How it does so is uncertain, as the Epidiolex label states, *"The precise mechanisms by which EPIDIOLEX exerts its anticonvulsant effect in humans are unknown. Cannabidiol does not appear to exert its anticonvulsant effects through interaction with cannabinoid receptors."* Whether CBD is converted to THC in the body is unclear, and despite it having no endocannabinoid activity, many people swear by its pharmacologic effects to reduce anxiety, chronic pain, and stress. Some even go so far as to say that CBD has anti-tumor and anti-cancer properties. Metabolically, CBD competitively inhibits a different set of enzymes: CYP3A4, CYP2B6, CYP2C9, CYP2D6, and CYP2E1. In all cases, CBD was a more potent inhibitor than THC, which would suggest that the DDI potential for CBD is higher than for THC [328].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The problem with assessing THC's DDI potential is the lack of clinical studies; you would be hard-pressed to find any dedicated clinical Phase 1 studies in patients or healthy volunteers assessing the DDI with THC. For decades, actual clinical research was limited because of government restrictions. Therefore, assessing THC's DDI potential must be done indirectly - THC inhibits this enzyme, this enzyme metabolizes this drug, ergo, THC could interact with this drug. An example of this is the reported interaction between THC and theophylline. There are no actual clinical studies demonstrating this interaction, but this is a DDI that is reported in many clinical DDI databases. The rationale is as follows: THC increases CYP1A2 expression, CYP1A2 metabolizes theophylline; therefore, theophylline effectiveness is decreased with long-term concomitant administration of THC. This prediction is not such a stretch, since there is a known interaction that smoking cigarettes decreases theophylline's effectiveness via CYP1A2 induction, but this cannot be considered a blanket statement. THC consumption can vary by daily intake, administration frequency, route of administration, and other factors, and it is unclear how these factors affect CYP1A2 expression and how they may modulate the potential interaction.

Does CBD Get Converted into THC?

There is considerable debate as to whether CBD is converted into THC in the body [329]; this is why in Figure 85 there is a "???" under the step from CBD to THC. It is well accepted that CBD can be converted in a test tube to THC under very acidic conditions and that this is a non-enzymatic process. Using simulated gastric fluid (pH ~1.2) in a test tube, the conversion can be demonstrated. However, in the body, the pH under which this spontaneous change occurs is too acidic, even under fasting conditions, for the conversion to occur. If CBD is converted into THC, then THC should be detected in the blood after administration of CBD. Oral administration of CBD has had mixed results – some studies detected THC in the blood, some have not. It has been argued that early CBD studies were flawed from faulty analytical methods that were not specific enough to detect THC, and that many of the formulations were far from pharmaceutical grade and were contaminated with THC to begin with. Even though the conversion of CBD to THC in the body is far from certain, and CBD appears to be pharmacologically inert, millions of people still take CBD to self-medicate and treat a variety of ailments.

THC and CBD Product Quality

THC and CBD products are not held to the same product quality standards as pharmaceutical drugs. Quality and potency can vary across manufacturers and even across plants within the same manufacturer. One analysis of CBD quality in 56 gummy products purchased online found that CBD concentrations differed by more than 10% from the stated packaging amount in 70% of the products tested. Further testing of gummies within products showed that gummy CBD concentrations varied by up to 27%. The authors of the paper concluded that "*CBD product packaging is not an accurate indication of the product constituents*" and that "*CBD products do not have reliable CBD concentrations-both relative to the dose listed on the label and among individual gummies within the same bottle*" [330]. State regulators are much slower to ensure consistent, safe product quality, so there is some 'buyer beware' aspect here.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

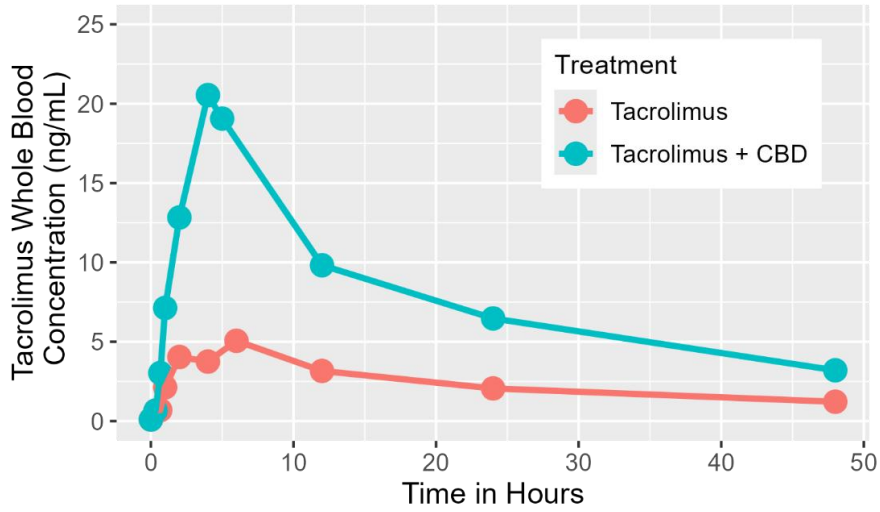


Figure 84: Tacrolimus whole blood concentrations in the absence and presence of CBD were administered to 15 healthy volunteers. Volunteers were administered 5 mg tacrolimus alone on Day 1. One week later, on Day 8, they received CBD 2.5 mg/kg once-daily for 4 days, which was increased to 2.5 mg/kg twice-daily on Day 12, and then increased again to 5 mg/kg CBD twice-daily on Day 15. On Day 24, they received 5 mg/kg CBD and 5 mg tacrolimus at the same time. Whole blood samples were collected on Day 1 and Day 24 for 48 hours and measured for tacrolimus. CBD increased tacrolimus concentrations about 4-fold. Figure modified from [331].

There are, however, many reported DDI studies done with CBD. CBD is a strong inhibitor of CYP2C19, a moderate inhibitor of CYP3A4, and a moderate inhibitor of CYP1A2. Drugs metabolized by any of these enzymes could have a potential DDI with CBD. One example is with clobazam, a drug used to treat epilepsy, which is metabolized by both CYP2C19 and CYP3A4. CBD has been shown to increase clobazam concentrations in patients by 3-fold [332]. Another example is tacrolimus, which is used to prevent rejection of solid-organ transplants. CYP3A4 mainly metabolizes tacrolimus. The first case report of an interaction between CBD and tacrolimus was in 2019, where tacrolimus concentrations increased 3-fold in a 32-year-old woman with refractory epilepsy who was taking tacrolimus after starting high-dose pharmaceutical-grade CBD (2,000 to 2,900 mg/day) [333]. A later dedicated Phase 1 DDI study between CBD and tacrolimus showed that CBD administration increased tacrolimus concentrations almost 4-fold (Figure 84) [331]. This interaction was significant enough for the manufacturer Astellas to change the Prograf (tacrolimus) label in 2023 to add the following sections:

1. In the Drug Interactions section of the Highlights, the following was added: *"Therapeutic drug monitoring and dose reduction for PROGRAF should be considered when PROGRAF is co-administered with cannabidiol".*
2. A new subsection was added to the Warnings and Precautions section of the label (Section 5.17): *"When cannabidiol and PROGRAF are co-administered, closely monitor for an increase in tacrolimus blood levels and for adverse reactions suggestive of tacrolimus toxicity. A dose reduction of PROGRAF should be considered as needed when PROGRAF is co-administered with cannabidiol".*
3. A new subsection was added to the Drug Interactions section of the label (Section 7.3): *"The blood levels of tacrolimus may increase upon concomitant use with cannabidiol. When cannabidiol and PROGRAF are co-administered, closely monitor for an increase in tacrolimus blood levels and for adverse reactions suggestive of tacrolimus toxicity. A dose reduction of PROGRAF should be considered as needed when PROGRAF is co-administered with cannabidiol".*

Further, in the Patient Information document, the CBD interaction was added to notify patients.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

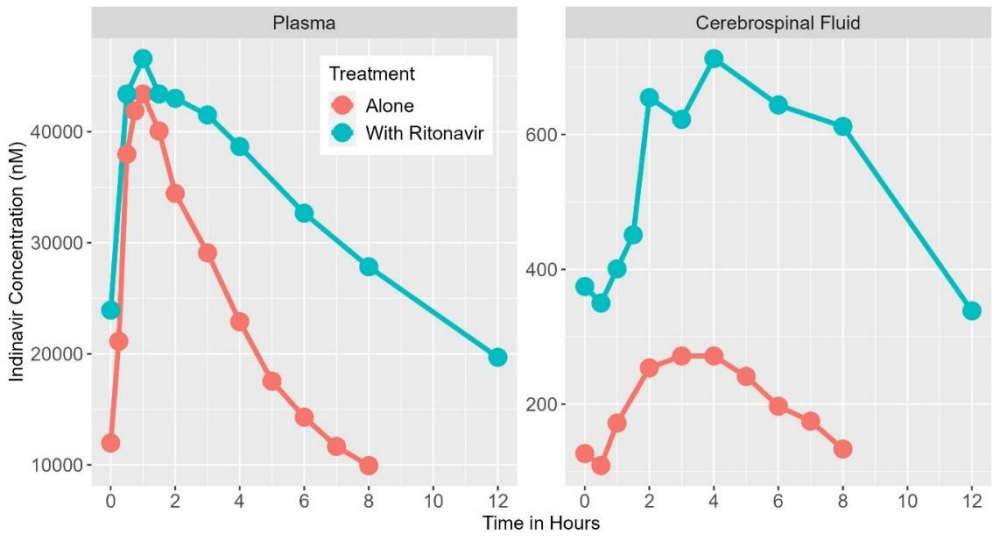


Figure 85: Ritonavir has been used to boost the efficacy of HIV drugs. Ritonavir is a potent irreversible inhibitor of CYP3A4, which inhibits the metabolism of indinavir, increasing brain and plasma indinavir concentrations in HIV patients. Figure redrawn from [334].

What has just been discussed has been THC's and CBD's potential to cause drug interactions with other drugs, i.e., their perpetrator effect. Other drugs have the potential to affect THC and CBD metabolism – they can be victims of a DDI. Strong CYP2C9 inhibitors, like fluconazole or amiodarone, and moderate inhibitors, like voriconazole, could inhibit THC metabolism and increase THC concentrations, potentially increasing its effect. But again, this is an indirect assumption, and there is no direct clinical evidence to support it. According to the Epidiolex label, strong CYP3A4 inhibitors, such as azole antifungals, and strong CYP2C19 inhibitors, such as fluvoxamine or fluoxetine, may increase CBD concentrations by 2-fold and potentially increase its effects. Long-term administration of CYP-inducing drugs, like rifampin, has been shown to decrease CBD concentrations, reducing its effect.

Not all Drug Interactions are Bad

Although the examples presented in this chapter have illustrated drug interactions that may result in harm or decreased efficacy for a patient, there are also drug interactions that increase the therapeutic effect of a drug. These interactions tend to potentiate the effect of one drug over the other. An example of a therapeutic drug-drug interaction is pharmacokinetic boosting or pharmacokinetic enhancement in patients with human immunodeficiency virus (HIV) using Norvir (ritonavir) or Tybost (cobicistat). The idea is that administration of drugs that inhibit the metabolism of an HIV medication will prolong the effect of the HIV medication and potentially overcome some forms of drug resistance. For example, patients with HIV may have neurologic sequelae as part of their illness. Protease inhibitors are standard therapy, but their penetration into the brain may be limited, leading to low protease concentrations. One such potent protease inhibitor is indinavir. Ritonavir is another protease inhibitor, but its use is limited because of its hyperlipidemic side effects. Ritonavir is, however, a potent, irreversible inhibitor of CYP3A4 and is often used to provide a short-term "boost" to other protease inhibitors. In HIV patients administered indinavir and low-dose ritonavir, both plasma and brain concentrations of indinavir were significantly increased compared to when ritonavir was not used (Figure 85) [334]. Unfortunately, outcomes were not reported in this study, but one can assume that ritonavir coadministration improved indinavir treatment outcomes. Today, fixed-dose combinations of HIV medications and boosters are the standard of care. For example, Kaletra is a

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

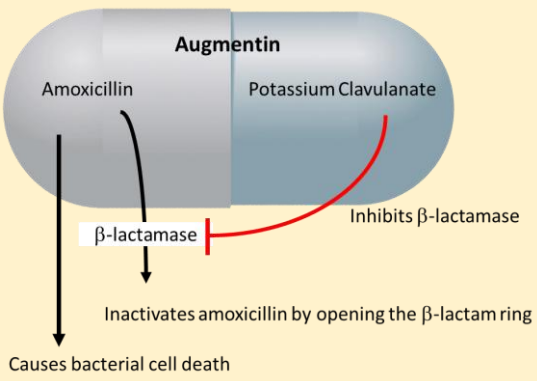
fixed-dose combination of lopinavir, a protease inhibitor, and ritonavir. Symtuza is a fixed-dose combination of four drugs: darunavir, cobicistat (the booster), emtricitabine, and tenofovir alafenamide. Fixed-dose combinations improve patient compliance because they only take a single pill, decrease HIV resistance rates, and have better tolerability than having to take many pills throughout the day.

Another example of a beneficial drug-drug interaction is the combination of DOPA and carbidopa (Sinemet) in patients with Parkinson's disease. Patients with Parkinson's disease have decreased dopamine in the substantia nigra of the brain, leading to tremors, dyskinesia, and difficulty walking. DOPA is a precursor for dopamine, crosses the blood-brain barrier, and is metabolized to dopamine in the brain. The movie Awakenings (1990), starring Robin Williams, dramatized the administration of DOPA to Parkinson's patients and the miracle it was. Patients who were essentially frozen and could not move were able to walk again. But as the movie showed, the effect was short-lived, and their symptoms returned. Carbidopa inhibits the peripheral metabolism of DOPA, allowing more DOPA to be taken up into the brain and thereby potentiating its effect. The combination of DOPA and carbidopa allows DOPA doses to be 75% lower than normal. A combination pill of DOPA and carbidopa was approved in 1975 and marketed under the brand name Sinemet.

The last example of a beneficial drug interaction is the combination of amoxicillin and potassium clavulanate, which is marketed as Augmentin. Amoxicillin is a penicillin derivative that is active against a wide range of bacteria. Amoxicillin alone has an extremely short half-life of ~ 1 hour owing to its rapid metabolism by β-lactamase, an enzyme produced by the bacteria that metabolizes the antibiotic, rendering it inactive; β-lactamase is a bacterial self-defense mechanism. Potassium clavulanate inhibits β-lactamase, thus restoring amoxicillin's effect. The combination also broadens the spectrum of bacteria against which amoxicillin is active, making it more effective against a wider range of bacteria.

Another Food-Drug Interaction

Popeye loved spinach. It's a good thing he wasn't taking warfarin. Leafy green vegetables are high in vitamin K and may interact with warfarin leading to increased risk of bleeding. These include asparagus, broccoli, Brussels sprouts, collard greens, kale, mustard greens, seaweed, spinach, Swiss chard, and turnip greens. Patients taking warfarin should maintain a consistent intake of vitamin K and avoid fluctuations from day to day or week to week.



Augmentin is an antibiotic that uses the pharmacokinetic interaction of two different drugs, amoxicillin and potassium clavulanate, for therapeutic benefit.

How Drug Interactions Are Studied Today

Before the late 1990s, drug interactions were studied empirically. Scientists looked at the drugs that might be co-prescribed with the new drug being developed, and then they conducted a study to see if a drug-drug interaction occurred between the two drugs. This was inefficient because it was not

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

uncommon for a company to submit more than 20 drug interaction studies to support a submission. And even with this approach, potential drug interactions could be missed. An example of this approach is lovastatin (Mevacor), the first statin for treating hypercholesterolemia, which was approved in 1987. Merck conducted drug interaction studies with the following drugs: cyclosporine, danazol, amiodarone, warfarin, colchicine, ranolazine, propranolol, digoxin, glipizide, chlorpropamide, and grapefruit juice.

Today, drug companies take a much more strategic approach. They determine how a drug is metabolized and whether any CYP enzyme predominates in its metabolism. They determine whether any transporter proteins transport their drug. They use recombinant CYP enzymes and isolated human hepatocytes to assess whether their drug inhibits CYP enzymes using known substrates. They also investigate whether their drug induces CYP enzymes. For example, suppose CYP3A4 and CYP2C19 predominantly metabolize a drug. The sponsor may then conduct clinical drug-drug interaction studies with potent inhibitors of these enzymes, like itraconazole for CYP3A4 and fluconazole for CYP2C19. In this case, the new drug is referred to as a potential 'victim' of a drug interaction. The drug causing the interaction is called the index 'perpetrator' because it has a known degree of enzyme inhibition against which other drugs can be compared. The idea behind this approach is that if no interaction occurs with the strongest metabolic inhibitor for a CYP enzyme, then no interaction will occur with weaker metabolic inhibitors. If an interaction occurs, the sponsor may then later conduct a study with a weaker inhibitor, depending on the magnitude of the interaction with the strong inhibitor. For example, if there were a 10-fold increase in drug concentrations with a strong inhibitor, the sponsor might consider conducting a study with a weaker inhibitor. However, if a strong inhibitor only doubles the drug concentration, then a study with a weaker inhibitor may not be warranted. Using this strategic approach reduces the number of needed drug interaction studies and increases the generalizability of the study results to other drugs.

If a new drug is an inhibitor or inducer of a specific enzyme or transporter, then studies can be conducted using probe substrates, which are prototypical, specific substrates for an enzyme of transporter, to estimate the magnitude of the effect in humans. In this case, we refer to the new drug as being a potential perpetrator of a drug-drug interaction, and the drug being affected as the index substrate because it has a well-established profile for comparison. These studies can be conducted with a single probe or as a cocktail approach, where all enzyme probe substrates are administered simultaneously and all probe drugs are then assayed in a single sample. One such cocktail is the "Pittsburgh Cocktail," which contains caffeine (for CYP 1A2), chlorzoxazone (for CYP 2E1), dapsone (for CYP 3A), debrisoquine (for CYP 2D6), and mephenytoin (for CYP 2C19) [335]. Of course, it is assumed that any drug in the cocktail will not affect the metabolism of any other drug in the cocktail. In this manner, the perpetrator effect of a new drug on several CYP enzymes can be determined in a single study. Other cocktails have been developed by studying different enzyme combinations. New drugs that increase the concentrations of a probe substrate more than 5-fold are referred to as a 'strong inhibitor,' 2-fold to 5-fold as a 'moderate inhibitor,' and 1.25-fold to 2-fold as a 'weak inhibitor.'

An alternative to conducting a dedicated clinical DDI study is to obtain drug concentrations in Phase 2 and Phase 3 clinical trials and then utilize population pharmacokinetics (presented in the chapter Principles of Pharmacokinetics). The idea is to compare drug concentrations in patients taking the concomitant medication of interest with those in patients not taking it. Using sophisticated mathematical modeling, the magnitude of the perpetrator effect on drug concentrations may be estimated. This is a useful approach often employed in the development of cancer drugs, as it is challenging to conduct dedicated DDI studies in these patients. A relatively recent development is the use of physiologically based pharmacokinetic (PBPK) models to predict the effects of enzyme or transporter inhibition. Scientists have become quite good at predicting metabolic drug interactions, such as those involving CYPs, and these predictions are often used in the label rather than conducting an actual clinical drug interaction study.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

How Drug Interactions are Reported on the Label and Medication Guide

In 2006, the FDA revised the structure of package inserts to enhance their readability. Under the current labeling structure, drug interactions are reported in the "Drug Interactions" or "Pharmacokinetics" sections of the label. Any major interactions are also reported in the highlights section, the first section a reader sees when the label is opened. Today, most drug interaction sections are written with respect to the mechanism of the interaction or around interactions (or lack of interactions) with particular drugs.

For example, Eliquis (apixaban) is a direct oral anticoagulant used to treat embolisms and prevent deep vein thrombosis. Apixaban is primarily metabolized by CYP3A4, with minor contributions from other cytochrome P450 enzymes, and is a substrate of P-glycoprotein (P-gp). The sponsor conducted 11 drug interaction studies at the time of filing, which were categorized around metabolism and/or transport, interaction with other anticoagulants or antiplatelet agents, or other commonly prescribed drugs. The following metabolic and/or transporter-related studies were completed:

- Digoxin and apixaban. Digoxin is a P-gp substrate and could compete with apixaban efflux from cells. No DDI was observed.
- Naproxen and apixaban. Naproxen is a P-glycoprotein (P-gp) inhibitor and may affect apixaban efflux from cells. No DDI was observed.
- Ketoconazole and apixaban. Ketoconazole is a strong CYP3A inhibitor and P-gp inhibitor. Ketoconazole could affect apixaban metabolism and efflux from cells. An ~2-fold increase in apixaban AUC and a 1.6-fold increase in Cmax were observed.
- Diltiazem and apixaban. Diltiazem is a moderate inhibitor of both CYP3A and P-gp. Diltiazem could affect apixaban metabolism and efflux from cells. A < 1.5-fold increase in apixaban AUC and Cmax was observed.
- Rifampin and apixaban. Rifampin is a strong inducer of both CYP3A and P-gp. Rifampin could affect apixaban metabolism and efflux from cells. Apixaban AUC and Cmax decreased by ~50% with rifampin cotreatment.

The following antiplatelet or anticoagulant interaction studies were conducted to examine their effect on apixaban pharmacodynamics (they were not necessarily thinking of a PK interaction during these studies):

- Aspirin and apixaban,
- Clopidogrel and apixaban,
- Aspirin and clopidogrel and apixaban,
- Naproxen and apixaban, and
- Enoxaparin and apixaban.

The following studies were conducted with other commonly prescribed medications:

- Atenolol and apixaban, and
- Famotidine and apixaban. Famotidine increases gastric pH and may affect apixaban absorption. Famotidine also inhibits the organic acid transporter and may affect apixaban absorption.

Hence, the drug interaction section in the package insert was written around CYP3A4 and P-gp. In Section 7 of the package insert, DRUG INTERACTION, the following is reported:

> *7 DRUG INTERACTION*
>
> *Apixaban is a substrate of both CYP3A4 and P-gp. Inhibitors of CYP3A4 and P-gp increase exposure to apixaban and increase the risk of bleeding. Inducers of CYP3A4 and P-gp decrease exposure to apixaban and increase the risk of stroke and other thromboembolic events.*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

7.1 Combined P-gp and Strong CYP3A4 Inhibitors

For patients receiving ELIQUIS 5 mg or 10 mg twice daily, the dose of ELIQUIS should be decreased by 50% when coadministered with drugs that are combined P-gp and strong CYP3A4 inhibitors (e.g., ketoconazole, itraconazole, ritonavir) [see Dosage and Administration (2.5) and Clinical Pharmacology (12.3)].

For patients receiving ELIQUIS at a dose of 2.5 mg twice daily, avoid coadministration with combined P-gp and strong CYP3A4 inhibitors [see Dosage and Administration (2.5) and Clinical Pharmacology (12.3)].

Clarithromycin

Although clarithromycin is a combined P-gp and strong CYP3A4 inhibitor, pharmacokinetic data suggest no dose adjustment is necessary with concomitant administration with ELIQUIS [see Clinical Pharmacology (12.3)].

7.2 Combined P-gp and Strong CYP3A4 Inducers

Avoid concomitant use of ELIQUIS with combined P-gp and strong CYP3A4 inducers (e.g., rifampin, carbamazepine, phenytoin, St. John's wort) because such drugs will decrease exposure to apixaban [see Clinical Pharmacology (12.3)].

Section 7 states that in patients taking apixaban 5 or 10 mg twice daily, the dose should be reduced by half when coadministered with a strong CYP3A and P-gp inhibitor. Eliquis is available in 2.5 and 5 mg tablets, so a 50% dose reduction is possible in these patients. However, no lower dose strength is available for patients taking 2.5 mg twice daily, so the label is more equivocal in these patients, stating that coadministration should be avoided. Section 7 also mentions clarithromycin, stating that no dose adjustment is necessary, despite its strong inhibition of CYP3A and P-gp.

In Section 12.3 of the package insert, Pharmacokinetics, the following is reported:

Drug Interaction Studies

In in vitro apixaban studies at concentrations significantly greater than therapeutic exposures, no inhibitory effect on the activity of CYP1A2, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2D6, CYP3A4/5, or CYP2C19, nor induction effect on the activity of CYP1A2, CYP2B6, or CYP3A4/5 were observed. Therefore, apixaban is not expected to alter the metabolic clearance of coadministered drugs that are metabolized by these enzymes. Apixaban is not a significant inhibitor of P-gp.

The effects of coadministered drugs on the pharmacokinetics of apixaban are summarized in Figure 2 [see also Warnings and Precautions (5.2) and Drug Interactions (7)]

In dedicated studies conducted in healthy subjects, famotidine, atenolol, prasugrel, and enoxaparin did not meaningfully alter the pharmacokinetics of apixaban.

In studies conducted in healthy subjects, apixaban did not meaningfully alter the pharmacokinetics of digoxin, naproxen, atenolol, prasugrel, or acetylsalicylic acid.

Section 12.3 states that apixaban does not affect the metabolism of other drugs, implying it does not act as a perpetrator of drug interactions. It then presents a plot showing the effect of inhibitors and inducers on apixaban C_{max} and AUC (Figure 86). Apixaban is considered the victim here. The type of plot shown in Section 12.3 is called a Forest Plot and is a common way to report pharmacokinetic data in the package insert. This plot is interpreted as follows. On the y-axis, a description of the treatment and the PK parameter being reported. Let's use the very first treatment in the plot: the effect of ketoconazole 400 mg on Cmax. The x-axis shows the Change Relative to the Reference. The Reference treatment is the C_{max} of apixaban alone. The Change Relative to Reference is approximately calculated by[8]:

[8] Technically, because pharmacokinetic parameters, like AUC and Cmax, are log-normally distributed, the Change

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

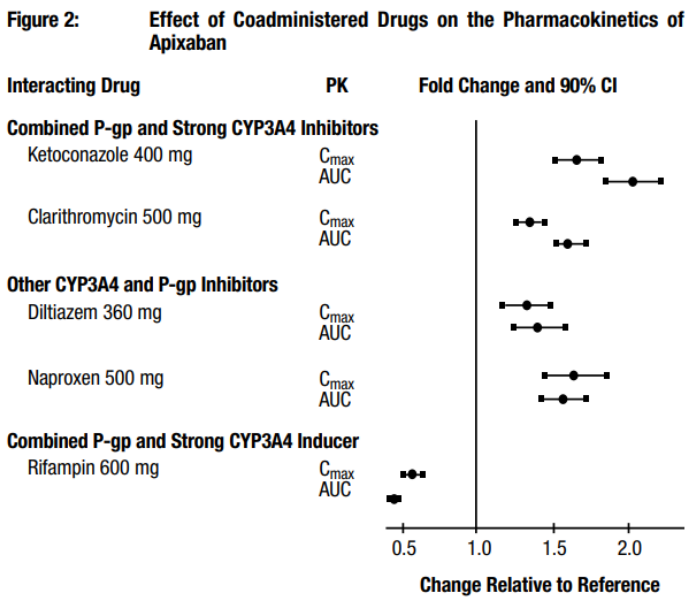


Figure 86: Forest plot shown on the Eliquis Package Insert in Section 12.3.

$$\text{Change Relative to Reference} = \frac{\text{apixaban Cmax when coadministered with ketoconazole}}{\text{Cmax of apixaban alone}}. \quad (38)$$

A value of '1' would indicate that ketoconazole does not affect the C_{max} of apixaban. A value of '2' would indicate that ketoconazole doubles the C_{max} of apixaban. A value of 0.5 indicates that ketoconazole decreases the C_{max} of apixaban by half. In clinical studies, ketoconazole 400 mg increased apixaban C_{max} from 140 to 225 ng/mL, a fold change of ~1.6 (=225/140)[9]. In the plot, this value is presented as a dot (●). Forest plots also present the imprecision in the estimate in the form of a confidence interval, shown as the bars around the dot ending in the square (■–– or ––■). In a sense, the confidence interval indicates the precision of the estimate. A large confidence interval indicates the estimate is not very precise. If the lower bound of the confidence interval includes the value of '1', then the estimate can be considered to be of no consequence, i.e., ketoconazole does not affect the Cmax of apixaban. In this case, the estimate appears to be fairly precise and does not contain '1'.

Particularly noteworthy drug interactions are reported in the highlights section of the package insert. This is the first part of the package insert, where the most important information is presented at a high level. In the Eliquis label, in the Drug Interaction highlights section, it states:

- Combined P-gp and strong CYP3A4 inhibitors increase blood levels of apixaban. Reduce ELIQUIS dose or avoid coadministration. (2.5, 7.1, 12.3)
- Simultaneous use of combined P-gp and strong CYP3A4 inducers reduces blood levels of apixaban: Avoid concomitant use. (7.2, 12.3)

The patient can quickly see that there is a problem with apixaban and strong CYP3A and P-gp

Relative to Reference is calculated as exp(Ln(apixaban Cmax when coadministered with ketoconazole) – Ln(Cmax of apixaban alone), where Ln() is the natural log function. For small numbers, however, this equation is approximately equal to Eq. (38).

[9] The raw values of 140 ng/mL and 225 ng/mL for C_{max} were not reported in the package insert, but were reported by the FDA in their review of the data submitted by the sponsor at the time of submission (https://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202155Orig1s000ClinPharmR.pdf).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

inhibitors; they can also refer to other sections for further details.

Particularly dangerous interactions can get a black-box warning, such as the black box warning for Lo Loestrin Fe, a combination oral contraceptive drug, which has a black-box warning against smoking:

> *WARNING: CIGARETTE SMOKING AND SERIOUS CARDIOVASCULAR EVENTS*
>
> *See Full Prescribing Information for complete boxed warning.*
>
> *Women over 35 years old who smoke should not use Lo Loestrin Fe*
>
> *Cigarette smoking increases the risk of serious cardiovascular events from combination oral contraceptive (COC)*

The Eliquis label points out some of the difficulties clinicians face when reading the package inserts. First, the clinician must know which drugs are strong CYP3A and P-gp inhibitors. There is also some language in Section 7 that is difficult to interpret. For patients taking 2.5 mg twice daily, it states that coadministration with a strong CYP3A and P-gp inhibitor should be avoided. What exactly does that mean? Is it OK to take together or not? Should the dosing regimen be altered, such as to 2.5 mg once daily? And why is clarithromycin called out specifically? Also, although Forest plots have been heralded as an easy way to present drug interaction data [336], they are not without their difficulties. How many people understand what fold-change is? Do they know that a 1.5-fold change equals a 50% increase? Also, how does the reader interpret the 90% confidence interval on this graph? Do they interpret it as the change that might happen in 90% of patients (the wrong interpretation), or that this is an estimate of the imprecision around the value of the reported fold-change (the correct interpretation)?

Other sources of information about possible drug interactions are the Medication Guide and the Patient Package Insert, which were mentioned earlier in the chapter on Regulatory Considerations. The Medication Guide is a patient-friendly document that outlines the risks and benefits of the medication. In contrast, the Patient Package Insert (PPI) is a more user-friendly package insert intended, ironically, despite its name, for healthcare professionals. Both documents are written by the manufacturer and approved by the FDA. Medication Guides must be distributed to patients. In both the Medication Guide and PPI, there is to be a section titled "What should I avoid while taking Drug-X?" It is in this section that any relevant drug interaction data should be disclosed. Interestingly, a review of the Eliquis Medication Guide does not mention any potential drug interactions; in fact, it does not even include a section on What to Avoid While Taking Eliquis. It's unclear why the Eliquis Medication Guide does not follow the mandated template, but it may be because it's grandfathered under old regulations.

As mentioned in the previous section, scientists have gotten quite good at predicting drug interactions using computer models. The benefit is that no studies are needed; patients or healthy volunteers are not unnecessarily exposed to a drug they do not need medically. This information is being incorporated into the package insert. For example, Padcev (enfortumab vedotin-ejfv) is an antibody-drug conjugate approved for the treatment of advanced bladder cancer. Drug interaction studies are difficult to conduct in cancer patients. Physiologically based PK (PBPK) models are a viable alternative to conducting actual studies. In Section 12.3 of the package insert, the results of the PBPK analyses are presented:

> *Drug Interaction Trials*
>
> *No clinical trials evaluating the drug-drug interaction potential of the ADC have been conducted.*
>
> *Physiologically Based Pharmacokinetic (PBPK) Modeling Predictions:*
>
> *Dual P-gp and Strong CYP3A4 Inhibitor: Concomitant use of PADCEV with ketoconazole (a*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

dual P-gp and strong CYP3A4 inhibitor) is predicted to increase unconjugated MMAE Cmax by 15% and AUC by 38%.

Dual P-gp and Strong CYP3A4 Inducer: Concomitant use of PADCEV with rifampin (a dual P-gp and strong CYP3A4 inducer) is predicted to decrease unconjugated MMAE Cmax by 28% and AUC by 53%.

Sensitive CYP3A substrates: Concomitant use of PADCEV is predicted not to affect exposure to midazolam (a sensitive CYP3A substrate).

The section clearly states that no human drug-interaction studies have been conducted. It then presents the results of the PBPK model: Padcev exposure is affected by CYP3A inhibitors and inducers, and does not affect other drugs metabolized by CYP3A.

Sometimes, drug interaction estimates are derived from population pharmacokinetic models. For example, Polivy (polatuzumab vedotin-piiq) is another antibody-drug conjugate, similar to Padcev, but used to treat specific types of leukemia. In Section 12.3 for the Polivy label, there is a section for predictions using both PBPK and population pharmacokinetics:

Drug Interaction Studies

No dedicated clinical drug-drug interaction studies with POLIVY in humans have been conducted.

Physiologically-Based Pharmacokinetic (PBPK) Modeling Predictions:

Strong CYP3A Inhibitor: Concomitant use of polatuzumab vedotin-piiq with ketoconazole (strong CYP3A inhibitor) is predicted to increase unconjugated MMAE AUC by 45%.

Strong CYP3A Inducer: Concomitant use of polatuzumab vedotin-piiq with rifampin (strong CYP3A inducer) is predicted to decrease unconjugated MMAE AUC by 63%.

Sensitive CYP3A Substrate: Concomitant use of polatuzumab vedotin-piiq is predicted not to affect exposure to midazolam (sensitive CYP3A substrate).

Population Pharmacokinetic (popPK) Modeling Predictions:

Bendamustine or Rituximab: No clinically significant differences in the pharmacokinetics of acMMAE or unconjugated MMAE are predicted when polatuzumab vedotin-piiq is used concomitantly with bendamustine or rituximab. Rituximab, Cyclophosphamide, Doxorubicin, or Prednisone (R-CHP):

No clinically significant differences in the pharmacokinetics of acMMAE or unconjugated MMAE are predicted when polatuzumab vedotin-piiq is used concomitantly with R-CHP.

Section 12.3 states that there was no apparent interaction with bendamustine or rituximab. Polivy is to be coadministered with R-CHP (a cancer drug cocktail of rituximab, cyclophosphamide, doxorubicin, and prednisone), which is the standard of care in patients with high-grade B-cell lymphoma. Polivy does not affect the pharmacokinetics of any of the components in this cocktail. When the results of a drug interaction are reported in a package insert, the FDA generally states the source of the information, whether it came from healthy volunteer studies, patient studies, PBPK models, PopPK models, or other computer simulations. However, a recent FDA guidance indicates that results from population pharmacokinetic studies will no longer be labelled as such.

Summary

Drug-drug interactions can occur at every step in the pharmacokinetic process: absorption, distribution, metabolism, and excretion. Interactions can also occur in the absence of changes in pharmacokinetics. It is beyond the scope of this chapter to present examples of each type of interaction; however, an attempt has been made to provide some representative and interesting examples. Drug interactions can also occur, not only between prescription drugs but also between

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

other types of drugs, such as drug-supplement and drug-food interactions.

Drug interactions are a fairly common problem, a significant number of which require hospitalization. In an ideal world, these interactions would be detected before the patient takes the combination. But in today's modern medical world, where patients have different doctors that might not communicate with each other, where patients use different pharmacies, perhaps a local pharmacy for some medications and a mail-order pharmacy for other medications, neither of which talks to each other, the likelihood of a drug interaction being missed can be high. Today, it falls on the patient to keep track of their medications and to *read the package insert*. Any questions can then be discussed with their doctor or their pharmacist.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Chapter 10

Personalized Medicine

There is a story in the press that keeps being told over and over. It goes something like this: Sometime in the near future, we may walk into our doctor's office, and they may take a blood sample, a biopsy, or another biological sample from us and send it to the lab for testing. Based on the test results, they will know not only which drug you will respond to, but also what dose you should be administered. They will then prescribe it. The idea that some biomarker, be that a measurement of our DNA, some substance or molecule in the blood, tissue, or tumor, or physiological measurements like blood pressure or hemoglobin A1c, may hold the secret to how we respond to drugs is both captivating and intriguing. This concept is called 'Personalized Medicine,' and some companies are banking their future on its success.

Traditionally, drugs are administered under a 'one size fits all' philosophy, with doses above or below the therapeutic range for patients who might need higher or lower doses. What we are talking about here is not only about finding the 'right dose,' but also getting the 'right drug' to the patient. This is the goal of personalized medicine, a topic that continues to intrigue and engage researchers and healthcare professionals alike. But who first proposed the concept of 'precision medicine'? Was it Paul Ehrlich at the turn of the 20th century, with the concept of his 'magic bullet'? Was it the Wall Street Journal in an article back in 1996? Was it US President Barack Obama who talked about his Precision Medicine Initiative in 2015? Or was it someone else? It's hard to say; each has been given some credit for it. Nevertheless, personalized medicine, synonymous with 'precision medicine,' is common parlance in drug development today. This chapter will explore how companion diagnostics are being used to implement the concept of personalized medicine, how more advanced techniques like -omics may be used to hyperpersonalize drugs, and how genetic differences can profoundly impact drug concentrations and responses.

Companion Diagnostics

Not-so-news flash: we are all different. Differences in absorption, distribution, metabolism, and excretion, as well as differences in genetics, environmental factors, receptor number, and receptor sensitivity, to name a few, all lead to differences in drug concentrations and response (Figure 87). As a result, some patients respond to drugs, while others do not. Some patients have side effects, and others do not. Some patients respond but experience side effects; others respond without side effects. Sometimes, the failure rate for a drug working as intended is quite high. It has been estimated that the failure rate is 38% for patients taking anti-depressive medications, 40% for asthma patients, 43% for diabetes medications, 50% for arthritis medications, and 70% for Alzheimer's medications [337]. For some medications, the failure rate exceeds the success rate.

There are many reasons for failure to respond to a drug:

- The patient is taking the wrong dose,
- The patient is not taking the drug properly,

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

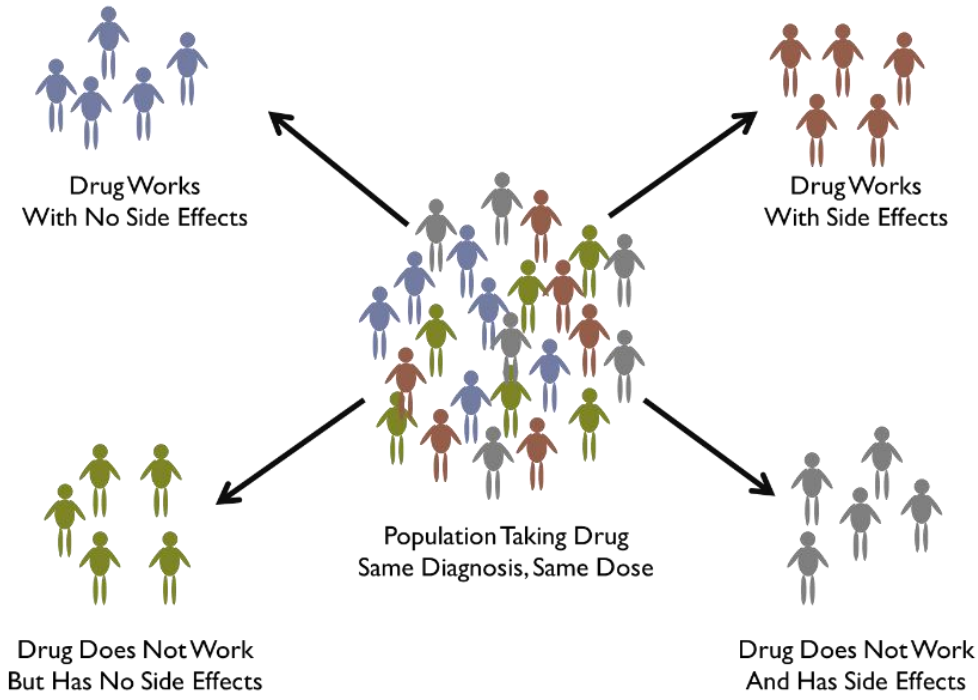
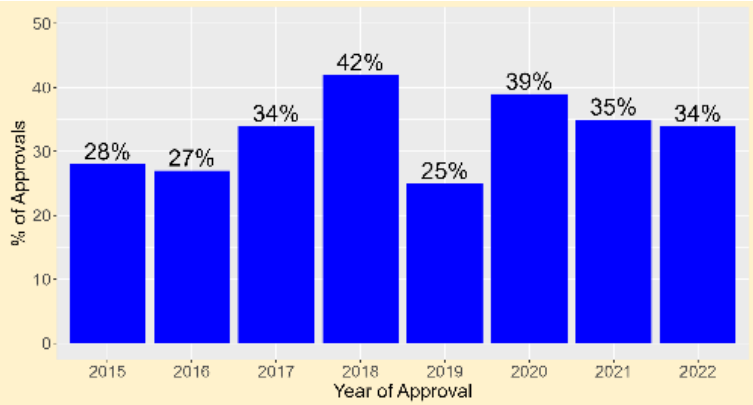


Figure 87: Not everyone responds the same to a drug.

- The patient has a different disease that the drug treats,
- The patient is taking the wrong drug,
- The patient has personal or genetic characteristics that affect the drug's PK and/or PD, and/or,
- The patient has concomitant diseases that may affect the drug's efficacy.

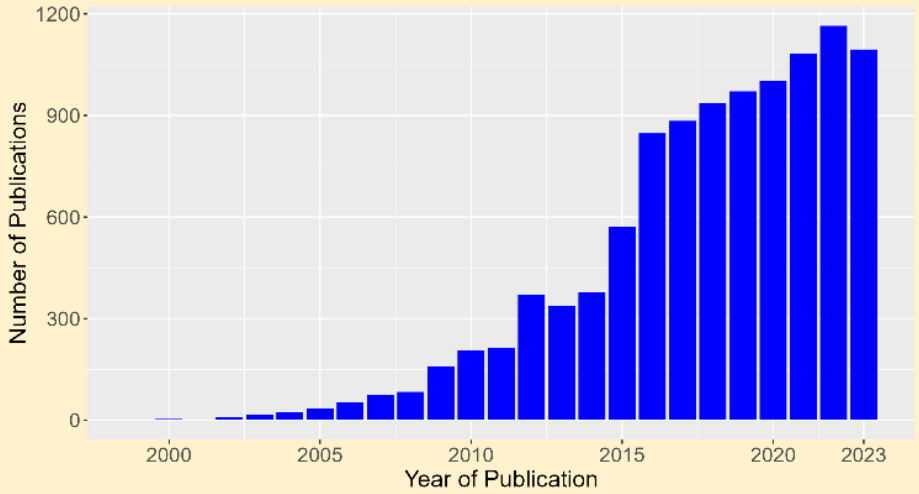
The outcome is unknown until the patient starts taking the medication. In precision medicine, patients are administered a laboratory diagnostic test called a companion diagnostic. Based on the test results, the decision is made whether to start the patient on the medication or to make dose modifications to an already prescribed medication.

Companion diagnostics are at the heart of personalized medicine, as the test is often developed alongside the drug. These tests must also receive FDA approval before use.



Between 2015 and 2022, 25% to 42% of Drug Approvals by the FDA were Accompanied by an Approval of a Companion Diagnostic [338]**.** In 2023, 38 of 52 approved drugs had some precision medicine component to their labels.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Number of publications with the words 'precision medicine' or 'personalized medicine' in their titles according to the National Library of Medicine's PubMed database. Publications were virtually nonexistent before 2000 but have exploded since then.

This means that they must accurately measure what they are supposed to measure with good precision and reproducibility. Companion diagnostics segment a population of interest based on the presence or absence of some genomic marker, demographic characteristic, or biomarker to ultimately improve treatment outcomes, reduce the incidence of some adverse event, or both.

Recent examples of drugs approved in 2023 that used an associated companion diagnostic are:

- ROS1 is a tyrosine kinase and is polymorphic in about 1% to 2% of the population. Augtyro (repotrecitinib) was approved for the treatment of patients with local or advanced metastatic non-small cell lung cancer who are ROS-1 positive in tumor biopsies.
- Cystic fibrosis transmembrane conductance regulator (CFTR) is the ion channel that controls the flow of chloride ions in and out of cells. Kalydeco (ivacaftor) is indicated for the treatment of cystic fibrosis (CF) in patients aged 1 month and older who have at least one mutation in the CFTR gene, based on clinical and/or in vitro assay data.
- Proteinuria is high protein levels in the urine and can result from kidney damage, inflammation, cancer, trauma, or intense exercise. Filspari (sparsentan) is indicated to decrease proteinuria in adults with primary immunoglobulin A nephropathy at risk of rapid disease progression. This product's use generally depends on whether the patient's urine protein-to-creatinine ratio exceeds ≥1.5 g/g; normal values are < 0.2.
- Welireg (belzutifan) is approved for treating von Hippel-Lindau disease, a rare disease that causes tumors in certain body parts. The tumors might not be cancerous, but they could be. Belzutifan is metabolized by CYP2C19 and uridine glucuronyltransferase (UDT2B27). The label recommends that "*Patients who are dual UGT2B17 and CYP2C19 poor metabolizers have higher belzutifan exposures, which may increase the incidence and severity of adverse reactions of WELIREG. Closely monitor for adverse reactions in patients who are dual UGT2B17 and CYP2C19 poor metabolizers.*"
- Familial hypercholesterolemia is a genetic condition passed from one generation to the next in which LDL (the bad cholesterol) concentrations can exceed 500 mg/dL (normal range: < 100 mg/dL). Homozygous familial hypercholesterolemia is when the patient inherits a defective gene from both parents. Angiopoietin-like 3 (ANGPTL3) is a protein that helps regulate LDL concentrations.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

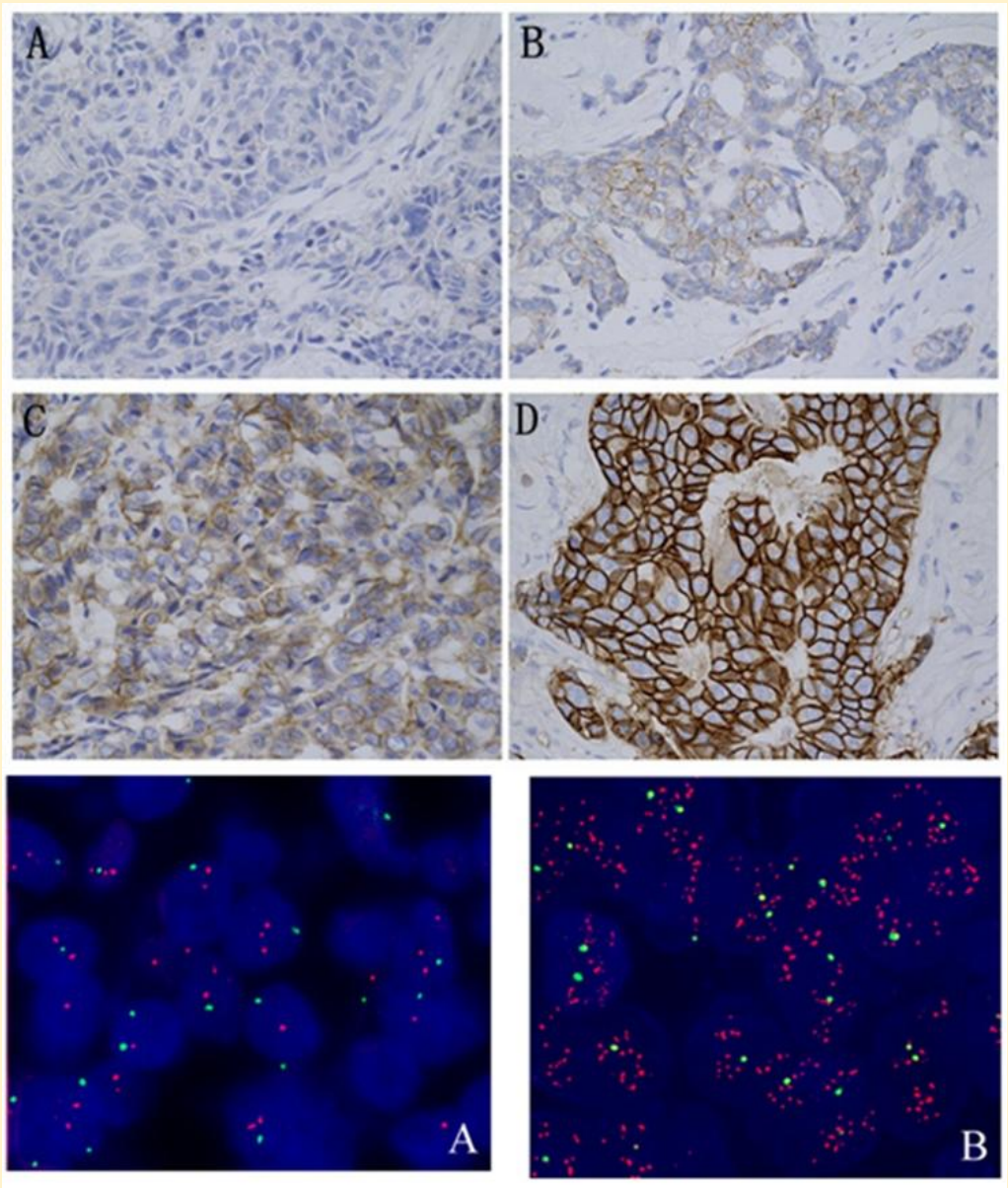


Figure 88: HER2 expression images. The upper and middle photos show HER2 protein expression by immunohistochemistry. Upper left: negative (score 0); upper right, negative (score 1+); middle left, equivocal (score 2+); middle right, positive (score 3+). The bottom row photos are of the HER2 expression by FISH. Bottom left, negative; bottom right, positive.

Figure kindly provided by Ji H, Xuan Q, Nanding A, Zhang H, Zhang Q. The Clinicopathologic and Prognostic Value of Altered Chromosome 17 Centromere Copy Number in HER2 Fish Equivocal Breast Carcinomas. PLoS One. 2015 Jul 10;10(7):e0132824. doi: 10.1371/journal.pone.0132824. PMID: 26161550; PMCID: PMC4498752.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Evkeeza (evinacumab-dgnb) is an ANGPTL3 inhibitor that is combined with other low-density lipoprotein-cholesterol (LDL-C) lowering therapies for the treatment of adult and pediatric patients, aged 5 years and older, with homozygous familial hypercholesterolemia.

The first targeted drug therapy to use a diagnostic test successfully was Herceptin (trastuzumab[10]) for treating metastatic breast cancer whose tumors overexpress a protein called HER2, which stands for human epidermal growth factor 2. HER2 is overexpressed in about 20% of breast cancers, making the patient HER2-positive, and is a growth factor that controls cell growth and division. Normally present at low levels, HER2+ cells overexpress HER2 on their cell membranes, leading to uncontrolled cell growth. Trastuzumab is a monoclonal antibody that binds to the HER2 receptor and inhibits its signaling, slowing cell growth. If the tumor does not overexpress HER2, trastuzumab will have little effect. Therefore, it is critical to determine patients' HER2 status before therapy. During the development of trastuzumab, patients were screened for HER2 status, and HER2-negative patients were excluded from the studies.

Herceptin showed that drugs requiring a companion diagnostic do not necessarily result in a loss of market share or revenue — quite the contrary. Sales of Herceptin exceeded $7 billion in 2015. Herceptin opened the door for other drugs with companion diagnostics.

There are many HER2 diagnostic tests, and considerable debate exists about which to use in practice. The gold standard, fluorescence in situ hybridization, or FISH for short, is 96% to 99% accurate at detecting a positive sample, given that the sample is truly positive (Figure 88). It is also 99% accurate at reporting a negative value when the sample is truly negative. The other common test is immunohistochemistry (IHC), where the biopsy sample is stained and examined under a microscope. ICH is not as accurate as FISH. It is only about 85% accurate at detecting a positive sample when it is truly positive and 94% accurate at reporting a negative value when it is truly negative [339]. The two tests report results differently. Whereas IHC reports a score of 0, 1+, 2+, or 3+, with 0 and 1 being 'negative,' 2 being borderline, and 3 being positive, FISH reports a simple 'positive' or 'negative' result.

There are many advantages to FISH, but one big disadvantage is cost, which is the primary reason it is not used exclusively. FISH is about 3-times as expensive as IHC, costing hundreds of dollars for the test. As a single agent, trastuzumab has been shown to prolong survival to 25 months in HER2-positive patients with newly diagnosed breast cancer with an overall response rate of 26% [340]. Subgroup analysis of breast cancer patients who were 3+ by IHC or FISH-positive had an overall response rate of 35% and 41%, respectively. Hence, patients with higher HER2 burden appeared to respond better than patients with lower HER2 burden.

Pharmacogenetics and Pharmacogenomics

Giving the right drug to a patient is just one part of the equation. The other part is giving the right dose. If a too-low dose is administered, the drug might not work; if a too-high dose is administered, it might not be safe. Choosing the right dose depends on a drug's pharmacokinetics and pharmacodynamics, which can be quite variable. Figure 89 shows the concentration-time profiles following a single dose of one of two formulations of praziquantel, a drug used to treat schistosomiasis, a disease caused by flatworms in the drinking water. It is commonly seen in Africa and parts of South America. Two formulations were studied: the marketed tablet and a new racemate

[10] A movie called Living Proof was made in 2008, starring Harry Connick Jr. as Dr. Dennis Slamon, the physician largely credited with discovering trastuzumab. Sadly, there was no movie for the inventor of the diagnostic test for HER2 status.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

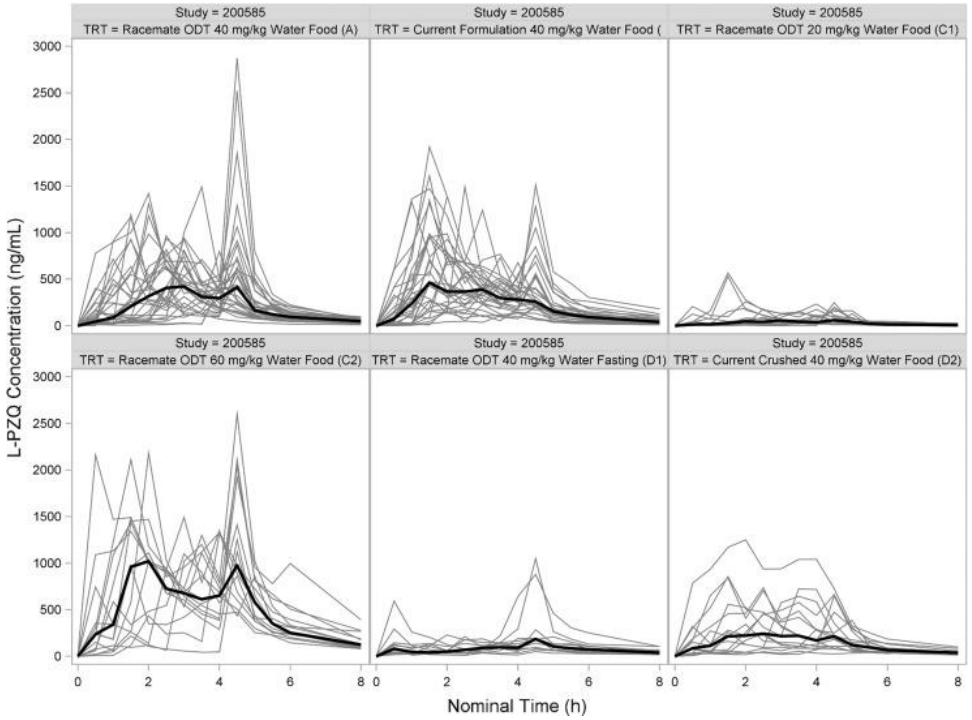


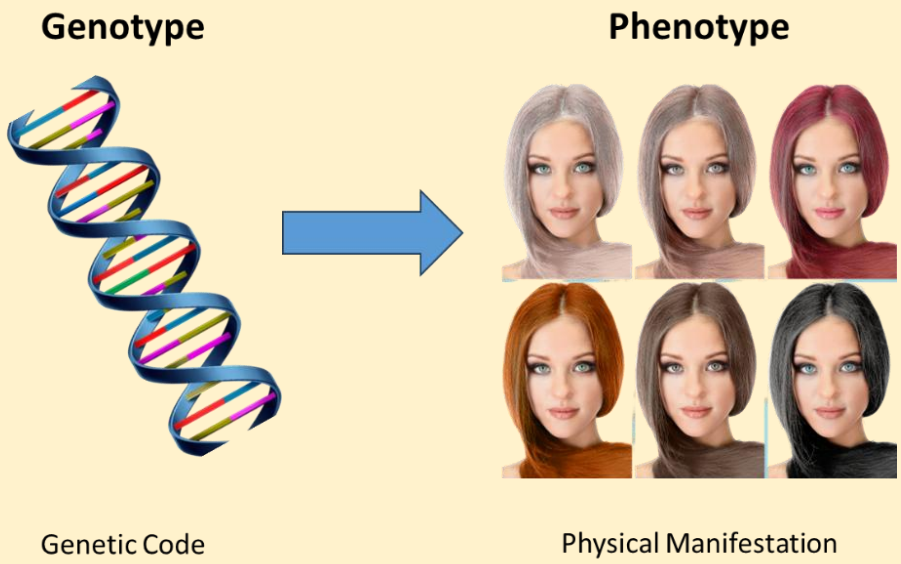
Figure 89: Plasma concentration-time profiles in adult subjects taking different formulations and doses of L-praziquantel illustrate the variability in concentrations across patients. The bold line is the mean. The letters in parentheses are treatment identifier codes. Figure reprinted from [341].

oral dispersion tablet. Two doses were studied: 20 and 40 mg/kg with water in either the fasting or fed state. The current market formulation, crushed into apple sauce, was also studied. Each gray line represents the profile of a single patient; the dark black line is the mean. Multiple peaks are observed in patients, rather than the single peak expected from an immediate-release formulation. Variability across patients is huge. Is this variability due to genetic differences? It is unlikely to be the sole reason, but it is most certainly a contributing factor.

Studying the effect of genetics on drug PK and/or PD has two names. Early studies in the latter half of the 20th century are referred to as 'pharmacogenetics' because those studies were often concerned with a single gene, e.g., what is the effect of CYP2D6 polymorphism on a drug's PK? Studies conducted today that use the patient's entire genome are referred to as 'pharmacogenomics' because we might be interested not just in one gene but in a set of genes that interact to affect a drug's PK and/or PD.

The role of genetics in affecting a drug's pharmacokinetics was first postulated in 1901 by Archibald Garrod [342], who would later write the seminal textbook *Inborn Errors of Metabolism*. Although the first textbook on the subject was published in 1962 [343], the field made little progress until the mid-1970s, when three groups independently discovered an interesting phenomenon. Michael Eichelbaum, Robert Smith, and Geoff Tucker, all European pharmacologists, noted that after administration of debrisoquine or sparteine, both drugs used to treat high blood pressure, a few patients began to experience blurred vision, headaches, and dizziness. The researchers collected urine from patients who received these drugs and found that, in patients with side effects, nearly all the drug was excreted unchanged into the urine; there were almost no metabolites. Most patients, however, did not experience side effects, and those who did had both the parent drug and multiple metabolites in their urine. This small group of patients who experienced side effects did not metabolize the drug to the same extent as those who did not [344].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Other terms used in genetics are genotype and phenotype. Genotype refers to the actual genetic code. For example, Arg389Gly refers to a change in arginine to glycine at position 389 in the amino acid chain. Phenotype is the physical manifestation of the genotype. A person can have a change in genotype with no change in phenotype. An example of different phenotypes is hair color. In pharmacogenetics, being designated a 'poor metabolizer' or 'extensive metabolizer' is an example of a phenotype associated with a genotype.

Further studies in larger groups of volunteers showed that about 90% of patients metabolized the drug, and 10% did not; the latter group was called 'poor metabolizers,' while those who did metabolize the drug normally were called 'extensive metabolizers.' Further studies showed that poor metabolizers had plasma drug concentrations 3- to 4-fold higher than those of extensive metabolizers [345]. Both debrisoquine and sparteine are metabolized predominantly by CYP2D6. This commonality was what led researchers to focus on that enzyme. Soon after that, a whole host of other drugs, like metoprolol, metoclopramide, and primaquine, were found to be polymorphically metabolized by CYP2D6. It was further discovered that the ratio of poor metabolizers to extensive metabolizers varies across different ethnic populations. For example, the original studies were done in Europeans, who are about 10% poor metabolizers. Asians, on the other hand, have only about 1% poor metabolizers, but Hawaiians or Pacific Islanders have about 22% poor metabolizers [346]. Later studies showed that the CYP2D6 metabolic effects were due to changes in the CYP2D6 gene.

Two additional terms are often used when discussing genetics: mutations and polymorphisms. Mutations are genetic differences that occur in less than 1% of the population. For example, if 99.9% of the population has guanine (G) at a particular point in their DNA, while 0.1% has a cytosine (C), this would be a mutation. Although mutations can be inherited from one of the parents, they are often spontaneous and result from DNA damage, such as chemical exposure, radiation exposure, smoking, or ultraviolet radiation from the sun. Conversely, polymorphisms are genetic sequence differences affecting more than 1% of the population. If 93% of the population had a G and 7% had a C, this would be a polymorphism.

An example of a polymorphism is blood type. In Caucasians, 44% are Type O, 42% are Type A, 10% are Type B, and 4% are Type AB. Polymorphic ratios are often different in different ethnicities; e.g., the incidence of Type O blood in Native Americans is 79%. Polymorphisms tend to exist naturally in the environment and are more related to evolution. They are passed from one generation to the next. Going back to the metabolism of drugs by CYP2D6, it was clear that the poor metabolism/extensive

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

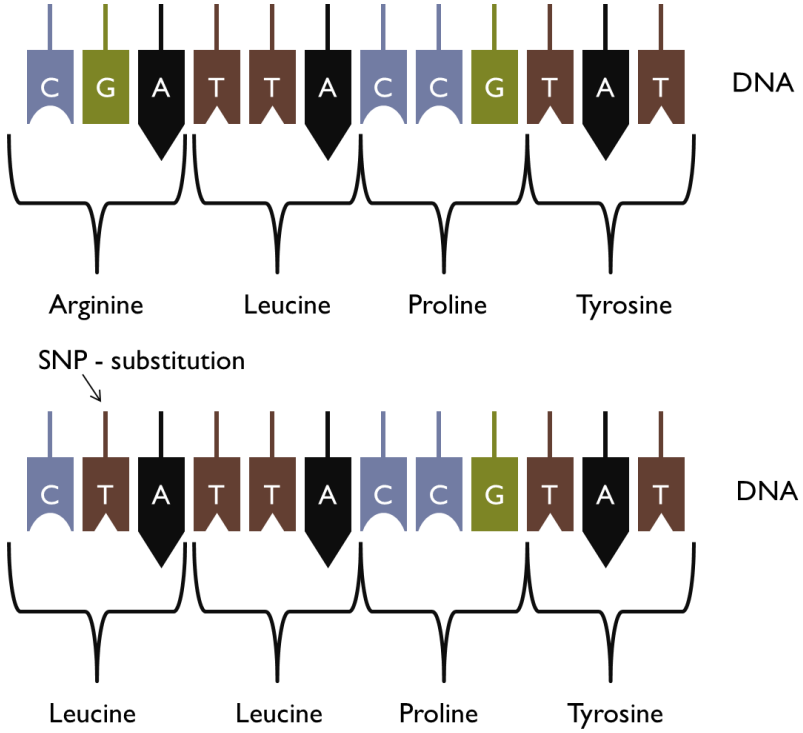
metabolism effects were due to polymorphisms in the population.

Mutations and polymorphisms can occur due to:

1. Changes (swaps) in a gene at the single-nucleotide level,
2. Deletions of nucleotides from a gene,
3. Insertion of additional nucleotides into a gene,
4. Or changes in the entire gene, such as multiplication or deletion.

Changes at the single-nucleotide level, called single-nucleotide polymorphisms (or SNPs for short), occur naturally at a rate of about 1 in every 300 base pairs in the DNA. Since there are about 3 billion nucleotide pairs in human DNA, there are over 10 million SNPs in a person's DNA. SNPs are what make us 'us.' They are what make us who we are. If we all had exactly the same DNA, we would look exactly the same.

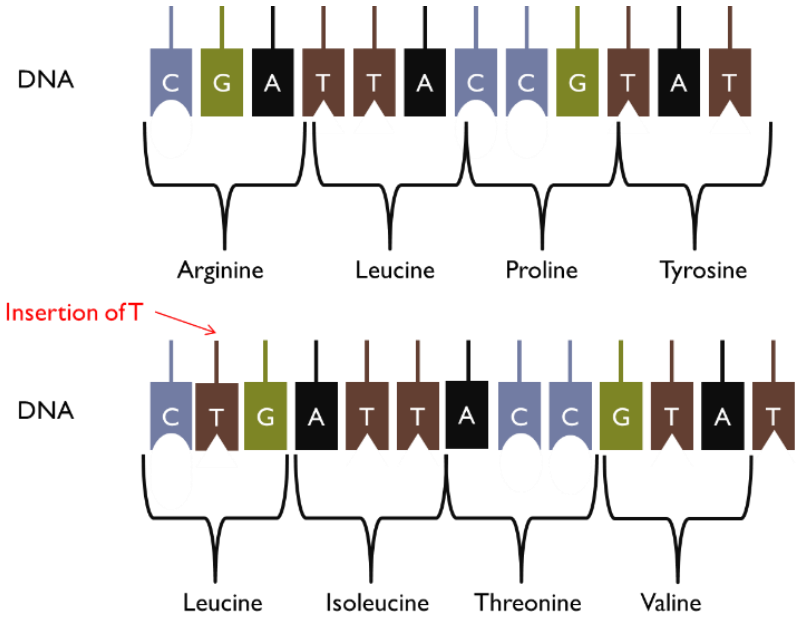
Recall that genes code for proteins. To be clear, SNPs in a gene do not always lead to a change in protein function. An example SNP would be the change from G to thymine (T) below:



DNA is transcribed into RNA, and RNA nucleotides are read in triplets, like a book, from left to right. Each triplet codes for a particular amino acid, and the sequence of triplets forms a sequence of amino acids that make up the protein. A single nucleotide change in DNA may result in a change in the amino acid synthesized. In this example, a change from G to T in DNA results in a change in the amino acid from arginine to leucine.

Nucleotide insertions and deletions are just what they sound like. A single nucleotide is inserted into the gene, or a single nucleotide is deleted from the gene. These changes are much more deleterious because the entire amino acid sequence can be altered, altering how it is read. For example, the insertion of a T between the first two nucleotides in our example results in an entirely different amino acid sequence:

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



A change in a single amino acid out of hundreds or thousands of amino acids in a protein may not seem like much, but it can have profound consequences. Sickle cell disease is caused by a single change in an amino acid in the gene that codes for hemoglobin. The amino acid glutamate is replaced by valine in position 6 of the beta-chain for hemoglobin due to an A to T transversion in the nucleotide sequence. Glutamate is a charged amino acid, but valine is not. Hence, valine makes the hemoglobin protein 'sticky' and clumps together into the characteristic sickles seen under a microscope.

An example of a single-nucleotide deletion leading to profound clinical consequences is cystic fibrosis, in which the deletion of a single nucleotide results in the loss of a critical amino acid in the CFTR

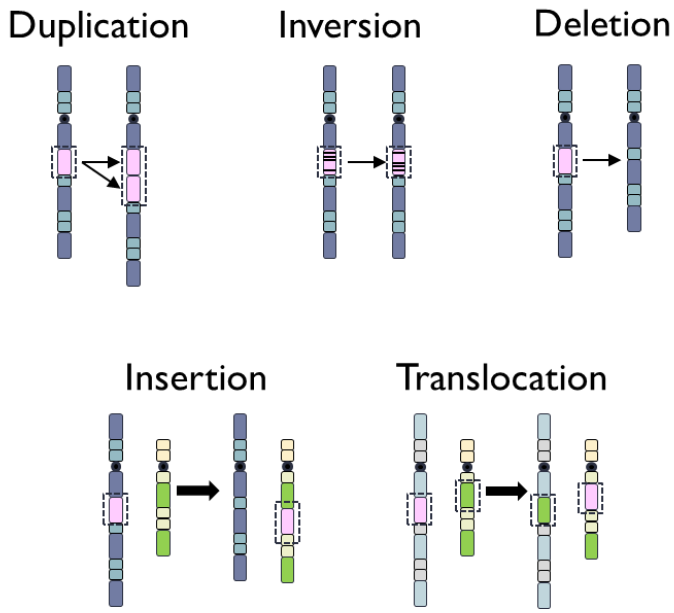
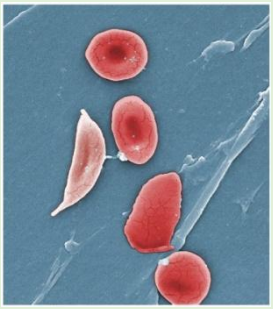


Figure 90: Mutational and polymorphic changes at the gene level.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

protein. This protein is critical for maintaining salt and water balance at cell membranes, such as those in the lungs.

Changes can also occur at the level of the gene (Figure 90). Genes can be duplicated, as seen in Charcot-Marie-Tooth disease type 1A, where multiple copies of several genes lead to excessive protein production, ultimately resulting in muscle weakness and movement difficulties. Genes can also be inverted, as seen in the case of hemophilia. Researchers found that the gene for hemophilia is the same in patients with hemophilia and those without, but that a hemophilia patient's gene was inverted on the chromosome, leading to improper reading of the DNA sequence and disruption of the clotting process. Genes can be deleted or inserted, and genes can be translocated. Gene translocation was discussed earlier in the book regarding bcr-Abl, a fusion gene resulting from the translocation of genes from chromosomes 9 and 22, creating the so-called Philadelphia chromosome, the causative gene for CML, and Gleevec, the first targeted molecular therapeutic targeting the tyrosine kinase bcr-Abl.



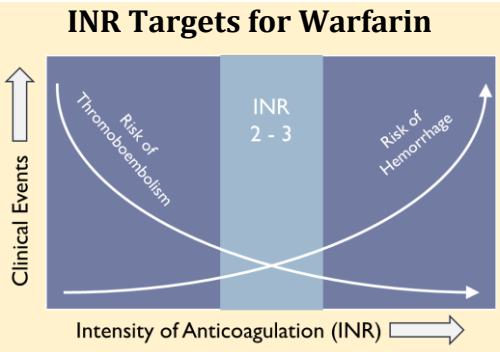
Blood smear from a patient with sickle cell disease.

Figure reproduced from Wikimedia Creative Commons.

Case Study in Pharmacogenetics – Drug Metabolism

CYP2D6 is not the only drug-metabolizing enzyme that shows genetic polymorphisms. After polymorphisms in CYP2D6 were discovered, scientists began examining other CYPs for polymorphisms, and several were identified. CYP1A2, CYP2B6, CYP2C8, CYP2C9, and CYP2C19 were all found to be polymorphic. CYP3A4 is polymorphic, but all of its polymorphs are either rare or have little change in phenotype. The allelic frequencies of many polymorphisms vary by ethnic group. This has been a common finding across polymorphisms in general – their allelic frequencies vary by ethnicity. Several Phase 2 drug-metabolizing enzymes were also found to be polymorphic, including N-acetyltransferase and uridine diphosphate glucuronosyltransferase. These polymorphisms can impact both the pharmacokinetics and pharmacodynamics of drugs.

As an example, consider Coumadin (warfarin), which was first developed as a rat poison in 1948 by Dr. Karl Paul Link, a chemist at the University of Wisconsin. The compound was isolated from the blood of a dying cow who had eaten moldy sweet clover hay and was later named 'warfarin,' after the group that funded the research: the Wisconsin Alumni Research Foundation (WARF). Since the 1950s, warfarin has been used as an anticoagulant and is currently used to prevent venous



INR, or International Normalized Ratio, is a standardized measure of blood clotting that reflects prothrombin time and is used to monitor warfarin anticoagulation. A normal INR is about 1 in people not taking anticoagulants, but the target is 2 to 3 in patients on anticoagulant therapy.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

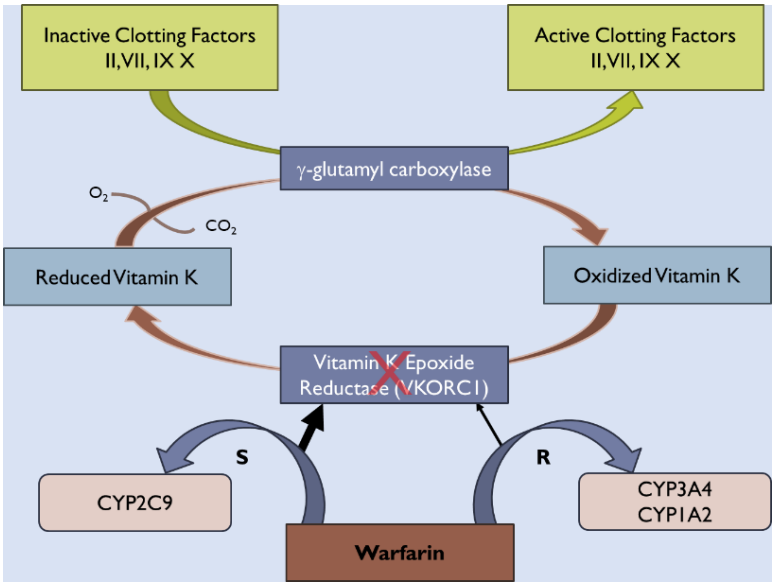


Figure 91: Overview of the mechanism of action of warfarin. Warfarin irreversibly inhibits the enzyme Vitamin K Epoxide Reductase (VKORC1), reducing synthesis of vitamin K–dependent clotting factors II, VII, IX, and X, thereby impairing blood coagulation. Figure redrawn from [347].

thrombosis, which is when a blood clot breaks off and blocks a vein; it is also used to reduce the risk of death, stroke, or recurrence following myocardial infarction. Despite the availability of newer, safer anticoagulants, warfarin sales still exceeded a half-billion dollars in 2022.

Warfarin is a difficult drug to dose correctly because it has a narrow therapeutic window and wide patient-to-patient variability. Below the therapeutic window, efficacy is lacking, while above it, there is an increased risk of bleeding events. For this reason, patient drug concentrations are titrated to the efficacious range. Many hospitals have warfarin or anticoagulation clinics where patients can have a blood draw to assess their INR (International Normalized Ratio). INR measures how quickly blood clots. The higher the number, the longer it takes for the blood to clot. Physicians want to slow blood clotting without prolonging the process. Physicians aim for an INR between 2 and 3. Generally, if the INR is below 2, the dose might be increased, whereas if it is above 3, the dose might be decreased. For first-time patients, finding the right dose can take a couple of weeks, as it's typically determined through trial and error. During this time, patients are at increased risk for adverse events and possible breakthrough emboli, which the drug is ironically being used to prevent.

Warfarin is a chiral molecule, with the S-enantiomer being 3- to 5-times more active than the R-enantiomer. S-warfarin is predominantly metabolized by CYP2C9, a polymorphic enzyme that converts it to inactive metabolites. CYP2C9 has more than 30 identified SNPs, but two are major: CYP2C9 *2 and *3. The wild-type CYP2C9*1 has about 80% frequency in Caucasians and close to 100% in African-Americans and Asians. *2 has about 70% of the activity of *1, is found in about 10% of the Caucasian population, and about <2% of African-Americans and Asians. *3 has about 20% of the activity of *1, is present in about 5% of the Caucasian population, and <2% of the African-American and Asian population. Therefore, patients who have *2 and *3 have decreased S-warfarin metabolism compared to wild-type patients and will require lower warfarin doses. But polymorphisms in metabolism alone do not account for all the variability in INR.

Warfarin's mechanism of action is complicated and is illustrated in Figure 91. In the coagulation process, gamma-glutamyl carboxylase is an enzyme that converts inactive clotting components into their active forms, which can then enter the clotting cascade. This step requires reduced vitamin K, which is then oxidized and inactivated. Oxidized vitamin K can be transformed back into its reduced,

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

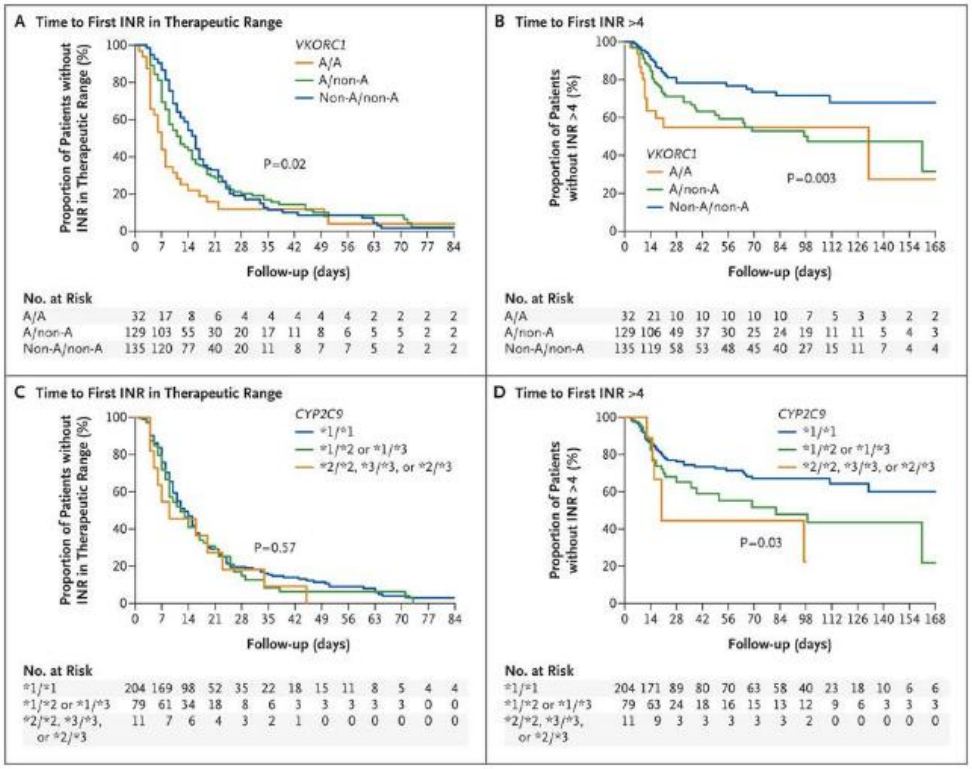


Figure 92: Kaplan-Meier plot showing the association between CYP2C9 and VKORC1 polymorphisms and INR outcomes as reported by [348], reprinted with permission from New England Medical Society.

active form by the enzyme vitamin K epoxide reductase (VKORC1). Warfarin irreversibly inhibits VKORC1, thereby preventing the conversion of oxidized vitamin K to reduced vitamin K and the conversion of inactive clotting factors to active clotting factors, ultimately preventing clotting. One of the reasons why warfarin is so difficult to dose is its irreversible mechanism of action. For many drugs that are competitive inhibitors of an enzyme, you can stop taking the drug, and the enzyme quickly reverts back to its normal state as the drug is cleared from the body. However, with warfarin, once warfarin irreversibly binds to VKORC1, it's permanently deactivated; you have to wait for the body to synthesize new VKORC1 to have activity levels return to normal, which can range from 4 to 12 hours.

VKORC1 is also polymorphic and contributes to INR variability. VKORC1 polymorphisms are classified differently from CYP2C9 and are classified by their haplotype instead of an SNP change. Haplotypes consist of genes that are inherited together, rather than as a single gene change. A common variant of VKORC1 is -1639 G>A, which manifests as an increase in warfarin sensitivity that results in reduced warfarin dose requirements. Like the CYPs, the frequency of VKORC1 polymorphisms differs by ethnicity. About 1 in 3 Caucasians carries the -1639 G>A polymorphism, whereas almost 90% of Asians do, which is why most Asians require lower warfarin doses than Caucasians. VKORC1 is generally classified as AA, AB, or BB for simplicity. Variant "A" is associated with low-dose warfarin, whereas variant "B" is associated with high-dose warfarin [349].

Does knowing a patient's CYP and VKORC1 status before starting warfarin therapy improve treatment outcomes? Schwarz and colleagues at Vanderbilt University set out to answer this question by enrolling 297 warfarin patients under standard treatment protocols and retrospectively

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

genotyping them for their CYP2C9 and VKORC1 polymorphisms. They then assessed whether there was a difference in the time to reach their first INR in the therapeutic range and the time to reach their first INR above 4 (Figure 92).

Patients with one or two VKORC1 haplotype A alleles quickly reached the therapeutic range, but also reached INR > 4 faster. Note that the non-A haplotype in the figure is similar to the "B" variant. CYP2C9 polymorphisms did not affect time to reach the therapeutic range, but *2 or *3 patients had an INR > 4 more often than patients with normal CYP2C9 function. After two weeks of therapy, CYP2C9 and VKORC1 polymorphisms also affected warfarin dose requirements. Patients with the VKORC1 A/A haplotype had lower warfarin requirements than non-A patients, as did patients with the CYP2C9*2 and *3 genotypes. While these results may suggest that genotyping is advantageous, the difficulty with this study and similar studies is that this is not how patients are treated with warfarin. Most patients do not have time to wait for genotyping before starting warfarin. They don't have the 5 to 10 days to get genotyped; they need to start warfarin therapy now.

The star allele nomenclature was created to make it easier for non-geneticists to name pharmacogenomic polymorphisms. This nomenclature is not used elsewhere in genetics, but only in pharmacogenomics. Instead of naming SNPs like g.27289G>C, the star allele system names them with a number and letter, e.g., *2a. *1 is usually denoted as the wild-type, the most common allelic frequency. *2, *3, *4, etc. denote polymorphisms that may lead to increased or decreased functionality. Because there are two alleles, the nomenclature uses two numbers, e.g., CYP2D6*5/*43; when one of the alleles is the wild-type, only the polymorphic allele is reported, e.g., CYP2D6*2. A new star allele nomenclature was introduced in 2017 because all the letters will eventually be used up; what happens after *3Z, for example? Under this new system, the letters have been replaced with .numbers; e.g., CYP2D6*2A has been changed to CYP2D6*2.001. A complete list of star alleles can be found at pharmvar.org.

To get patients into the therapeutic INR range more quickly, a variety of dosing algorithms have been proposed, more than 400 to date [350]. The algorithms consider not only CYP2C9 and VKORC1 status but also internal factors such as age, weight, and sex, as well as external factors such as smoking status and drug interactions. Unfortunately, most of these algorithms were created using small sample sizes, are not externally validated (compared against data not used to build the model), and have sufficient bias that, in one group's opinion, makes them not very useful for clinical use [350].

However, one algorithm of particular note is the International Warfarin Pharmacogenetics Consortium (IWPC), which does not suffer from these limitations. This study used 4053 patients and a validation dataset of 1009 patients to assess its performance [351]. The IWPC developed two algorithms: one based on genetic criteria and one on clinical criteria alone, without genetic information. The latter was far more predictive of the dose needed to be within the therapeutic range than the clinical prediction model.

The genetic model uses the following predictors: age, height, weight, VKORC1 genotype, CYP2C9 genotype, race, whether the patient is taking an enzyme inducer, and whether the patient is taking amiodarone. As an example, a Caucasian, 60-year-old male weighing 70 kg and being 5'10" tall (175 cm) with G/G VKORC1, *1 for CYP2C9, and not taking any concomitant medications, was predicted that their weekly warfarin starting dose should be 42 mg or 6 mg/day. Under this model, patients who are VKORC1 A/A require ~70% higher doses than G/G patients, and patients who are CYP2C9*1/*2 require almost half the dose as wild-type patients. Patients who self-identify as 'Black,' which is not the case in this example, have their dose reduced by 75%. With the genetic algorithm, 35% of dose predictions fell within 20% of the actual dose requirement, resulting in significantly fewer dose overestimations than with the clinical prediction model or a simple fixed-dose strategy. An online calculator using this algorithm can be found at WarfarinDosing.

Drug labels are slow to change. Warfarin has been around since the 1950s, and the genetic effects of polymorphisms on warfarin response have been known since the 1990s. However, it took until

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Another example of the inclusion of CYP polymorphism information in the label was the 2014 approval of Cerdelga (eliglustat) for the treatment of Gaucher disease. CYP2D6 predominantly metabolizes Eliglustat. Beyond poor metabolizers and extensive metabolizers, it is now known that there are other groups: intermediate metabolizers, whose activity lies between that of a poor metabolizer and an extensive metabolizer, and ultrarapid metabolizers, who have multiple copies of the CYP2D6 gene and produce much more CYP2D6 enzyme than the wild type. When ultra-rapid metabolizers are administered eliglustat, they metabolize the drug very rapidly, so these patients may not have sufficient drug concentrations to have an effect. Hence, administration of eliglustat in ultra-rapid metabolizers is not recommended. The package insert reads:

> *"CERDELGA is a glucosylceramide synthase inhibitor indicated for the long-term treatment of adult patients with Gaucher disease type 1 who are CYP2D6 extensive metabolizers (EMs), intermediate metabolizers (IMs), or poor metabolizers (PMs) as detected by an FDA-cleared test. CYP2D6 ultra-rapid metabolizers may not achieve adequate concentrations of CERDELGA to achieve a therapeutic effect."*

2007 for the warfarin label to be updated to include consideration of genetic testing in physicians' dose recommendations. The label was updated to:

> *"The lower initiation doses should be considered for patients with certain genetic variations in CYP2C9 and VKORC1 enzymes".*

In 2011, the label was updated again, including two proposed dosing regimens: one for patients with known genomic status and one for patients with unknown genomic status. For patients with unknown genomic status, the initial dose is 2 to 5 mg per day, with close INR monitoring and dose titration based on INR response. Maintenance doses usually range from 2 to 10 mg per day. For patients whose VKORC1 and CYP2C9 status are known, the following dose table is available:

Table 19

The FDA Drug Label for Coumadin
Warfarin Dose by VKORC1 Genotype

VKORC1	CYP2C9					
	***1/*1**	***1/*2**	***1/*3**	***2/*2**	***2/*3**	***3/*3**
GG	5–7 mg	5–7 mg	3–4 mg	3–4 mg	3–4 mg	0.5–2 mg
AG	5–7mg	3–4 mg	3–4 mg	3–4 mg	0.5–2 mg	0.5–2 mg
AA	3–4 mg	3–4 mg	0.5–2 mg	0.5–2 mg	0.5–2 mg	0.5–2 mg

Reference: [352]

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

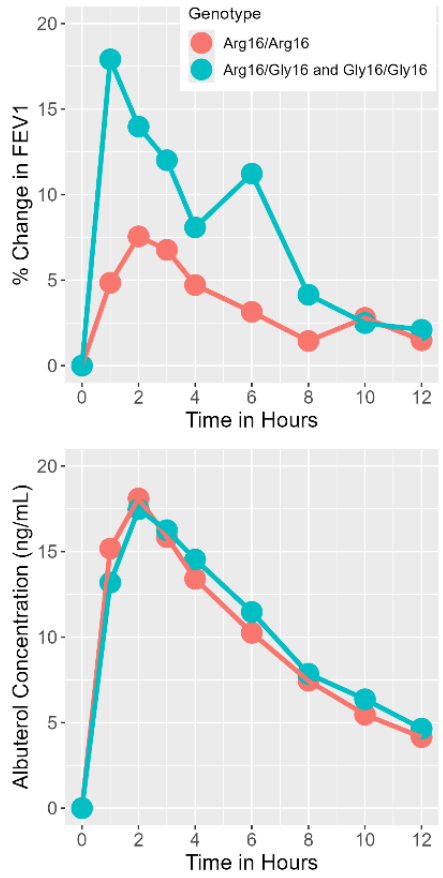


Figure 93: FEV1 and albuterol concentrations in 16 asthmatic patients administered a single dose of albuterol. Figure reprinted from [353].

Today, warfarin represents a label incorporating pharmacogenomic information to improve dosing for this difficult-to-use drug.

Study in Pharmacogenetics – Pharmacodynamics

Not only are many drug-metabolizing enzymes polymorphic, but it should be no surprise that receptors can also be polymorphic. These polymorphisms can increase or decrease an individual's sensitivity towards a drug. An example of this is seen with albuterol, a drug used to treat lung conditions like asthma. Albuterol is a bronchodilator that interacts with β2–β2-adrenergic receptors in the lungs and increases airflow into and out of the lungs. Albuterol's effect can be measured by the amount of air forced out of the lungs over 1 minute, or FEV1 (forced expiratory volume in 1 minute) for short.

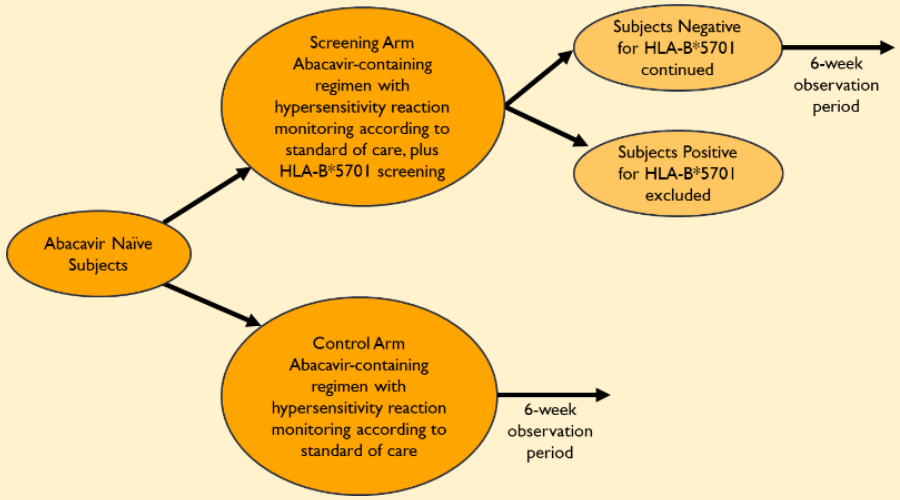
Several coding mutations have been identified in the β_2-receptor, resulting in receptors with different allelic frequencies and sensitivities. Of interest are the homozygous Arg16/Arg16, heterozygous Arg16/Gly16, and homozygous Gly16/Gly16 alleles. Gly16 is associated with increased down-regulation and reduced receptor sensitivity in cell culture. In one study, 16 patients with moderate asthma were administered a single dose of albuterol, and their FEV1 was monitored for 12 hours after dosing. Patients were genotyped for β2-receptor polymorphisms. Patients who were homozygous for Arg16 showed a greater and more rapid increase in FEV1 compared to patients who were heterozygous for Arg16 or homozygous Gly16, even though their albuterol plasma concentrations were the same (Figure 93) [353].

Skin Patch Test for Abacavir. With this test, patients are administered small amounts of abacavir in varying amounts. After a period of time, they are assessed for a skin rash, which indicates potential hypersensitivity to the drug if it is administered orally. This patient was tested after showing clinical signs of hypersensitivity: rash, fever, and gastrointestinal pain.

The figure was kindly provided by Margarida Goncalo, MD, PhD. [354].

Beta-receptors are not unique in their polymorphisms. Other receptors have polymorphisms that affect their response to drug treatment. Polymorphisms in opioid receptors modify pain tolerance and alter the pain relief response to non-steroidal anti-inflammatory drugs and fentanyl, as well as the propensity to abuse drugs [355]. Polymorphisms in HMG-CoA reductase, the enzyme target for statins, affect statin efficacy [356]. Polymorphisms affect the response of tyrosine kinase inhibitors in the treatment of cancer [357]. More on this point will be discussed later in this chapter on Personalized Medicine.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Study design for PREDICT-1

In the albuterol example, a polymorphism affected the drug's efficacy. Polymorphisms can also affect a drug's safety. A classic example is abacavir (Ziagen), which was approved in 1998 for the treatment of HIV infection. At the time of its approval, abacavir was given a black box warning on its label for a potentially fatal hypersensitivity reaction that occurred in about 5% of patients. This reaction was multisystemic, leading to fever, skin rash, fatigue, diarrhea, nausea, and vomiting, and started about six weeks after initiation of therapy. In some cases, anaphylaxis, liver failure, renal failure, hypotension, and death occurred. Treatment was to be stopped immediately and permanently discontinued. GlaxoSmithKline (GSK) was in a bind because no company wants any patient to experience an adverse reaction of this type. Still, they didn't have a reliable method to predict which patients might be at risk, aside from a skin patch test used by some investigators, later deemed unreliable by the FDA.

This led GSK to retrospectively re-consent patients for genomic analysis of samples from the clinical trials that approved abacavir. A wide-screening approach was used based on the genomic status of drug-metabolizing enzymes and human leukocyte antigens (HLA), proteins found on the surface of most cells that identify cells as "self." Drug-metabolizing enzymes were screened because of the known association between polymorphisms and drug effects. In contrast, HLAs were selected because there are hundreds of HLA proteins, allowing each person's immune system to react to a wide range of foreign invaders. Specific HLA alleles had previously been linked with many idiosyncratic adverse drug reactions associated with the immune system, as seen with abacavir. No association with drug-metabolizing enzymes was found, but a strong association between HLA-B*5701 positivity and the occurrence of hypersensitivity reaction was observed [66]. The association wasn't perfect – it could reliably predict hypersensitivity reactions about half the time. But half of the time, it was better than zero. Further studies with other datasets by other investigators confirmed the association. Also, more complicated methods, using what was at the time state-of-the-art, failed to identify any other marker that could predict hypersensitivity better than HLA-B*5701.

But this raised the question: Does knowing whether a person is positive for HLA-B*5701 decrease the risk of developing clinically significant hypersensitivity? Also, HLA-B, like other alleles, varies by ethnicity. Is there a difference in the ability of HLA-B*5701 to predict hypersensitivity in different populations? To address these questions, GSK did a prospective clinical trial called PREDICT-1 (Prospective Randomized Evaluation of DNA Screening in a Clinical Trial, NCT00340080). The study was designed to assess whether prospective screening and exclusion of patients positive for HLA-B*57001 resulted in a lower rate of hypersensitivity reactions than the control group, which was not genotyped.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

In this study, patients were randomized to a screening and a control arm. In the screening arm, patients were screened for HLA-B*5701. Positive HLA-B*5701 patients (~6% of the study population) were excluded from the study. Negative patients and control patients were then treated with abacavir for a 6-week observation period. The results showed that the hypersensitivity rate in the negative screened group was more than 50% lower than the control group: 7.8% in the control group compared to 3.4% in the negative HLA-B*5701 patients [358].

In 2008, the Ziagen label was updated to include this statement in the black box section:

> *"Patients who carry the HLA-B*5701 allele are at high risk for experiencing a hypersensitivity reaction to abacavir."*

The label also includes a recommendation for screening:

> *"Screening for carriage of the HLA-B*5701 allele is recommended before initiating treatment with abacavir. Screening is also recommended before reinitiation of abacavir in patients of unknown HLA-B*5701 status who have previously tolerated abacavir".*

But not by patch screening:

> *"Skin patch testing is used as a research tool and should not be used to aid in the clinical diagnosis of abacavir hypersensitivity".*

After the results of PREDICT-1, clinical testing laboratories began offering HLA-B*5701 genomic tests for patients. In 2015, the label language around screening became stronger, changing from just a recommendation to 'should be done':

> *"All patients should be screened for the HLA-B*5701 allele prior to initiating therapy with ZIAGEN or reinitiation of therapy with ZIAGEN, unless patients have a previously documented HLA-B*5701 allele assessment".*

Race-Based Drugs

With the idea of drugs tailored to an individual, an interesting drug is isosorbide and hydralazine in a fixed combination pill, which was approved as BiDil in 2005. BiDil was the first drug approved for the treatment of heart failure ***in Black patients***. The label for BiDil reads:

> *"BiDil is indicated for the treatment of heart failure as an adjunct to standard therapy in self-identified black patients to improve survival, to prolong time to hospitalization for heart failure, and to improve patient-reported functional status."*

Although the drug is approved for self-identified Black patients, it didn't start that way. Pivotal studies comparing BiDil to a placebo showed no difference in mortality. Subgroup analysis, however, suggested that Blacks may have had lower mortality than Whites. In a second study using the active comparator enalapril, subgroup analysis indicated that BiDil was inferior to enalapril, but only in the White group; Blacks had no difference in mortality compared to enalapril. This led to a third placebo-controlled study in self-identified Black patients. BiDil had lower mortality and fewer hospitalizations than placebo. Based on these results, BiDil became the first race-based drug.

This is where the label remains today. The screening program was a huge success. Recent estimates using meta-analysis put the hypersensitivity reaction rate at around 1.3% or less [359]. Furthermore, cost analyses indicate that the money saved from screening far outweighs the hospitalization costs should the patient have a hypersensitivity reaction to abacavir, leading many countries to adopt screening as the standard of care [360]. In what could have been a potentially drug-killing adverse event, GSK's use of a screening test saved the drug, and in 2023, more than 20 years after approval, abacavir still had sales exceeding a billion dollars a year worldwide.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Why Isn't Pharmacogenetics or Pharmacogenomics Used More Frequently in Clinical Practice?

It's just a matter of fact that pharmacogenetics and pharmacogenomics are not frequently used in clinical practice, and unlike companion diagnostics, are not used in drug development. It is rare for a newly approved drug to have a dosing regimen based on a patient's genetic information, as with Cerdelga (eliglustat). But why is this the case? If this information can reduce variability in drug concentrations, why isn't it used more often? Several reasons come to mind.

First, for many drugs, while genetics is important, it is not a major contributor to overall variability in pharmacokinetics. An oft-cited estimate puts 20% to 95% of the variability in a drug's response attributable to differences in genetics [361]. Genotyping a patient is also expensive and rarely covered by insurance, so why develop a dosing algorithm that may reduce variability marginally and require the patient to pay for the test needed to use it? Further, for most drugs the trial and error necessary to find the right dose is a patient is not an immediate need. For these drugs, it's okay to take a few weeks to find the right dose. Drugs that have genetic information in their label tend to be drugs that have an immediate need to find an optimal dose, like cancer drugs.

Lastly, it's really hard to establish the correlations needed to identify genetic markers that correlate with a pharmacodynamic effect. It's easier for SNP mutations like CYP2D6 polymorphisms to identify a difference, but more complicated ones, like the Abacavir example, take years and require thousands of patients. As a result, what you tend to see happen is that academic research centers do pharmacogenetic correlates after the drug is approved, and then, after years of research and petitioning, that information might make it into the drug label.

The Future of Personalized Medicine

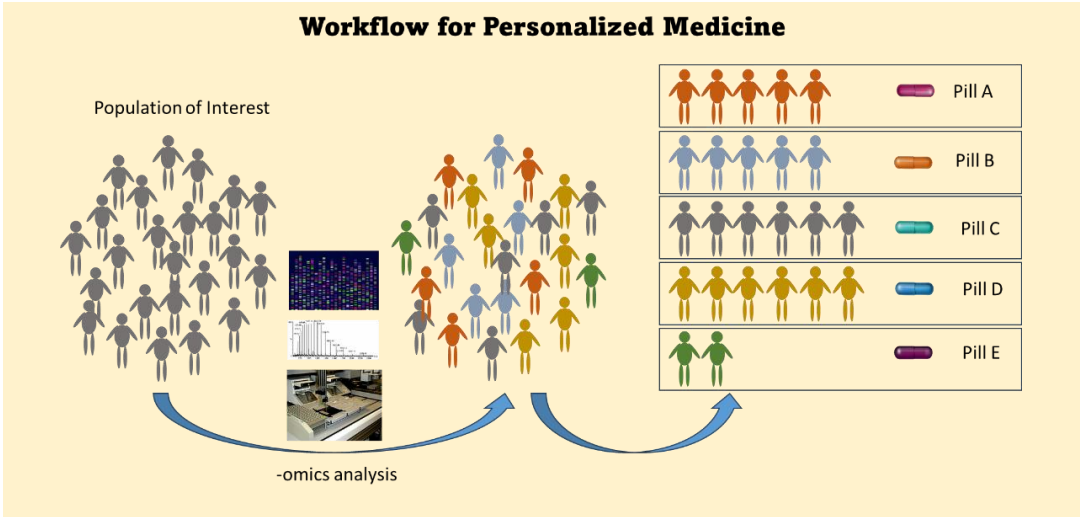
The future of personalized medicine is like companion diagnostics, but on steroids. Although the idea of precision medicine has been around since Hippocrates, success stories, such as imatinib, abacavir, and trastuzumab, began around the 1990s. Furthermore, the Human Genome Project, which was underway around the same time, promised that understanding the entire human genome would lead to breakthroughs in disease treatment. This sparked people's imagination that we could genetically screen an individual's entire DNA and then use that information to inform drug selection and dosing.

Instead of screening for a single molecular marker that might improve treatment outcomes or the safety profile, the idea is to use -omics technologies to better understand the patient's genetics and biochemistry. 'Omics' is a Latin word for mass or many and represents different bioanalytical techniques that measure many different endpoints in the blood, plasma, urine, tissue, or organ. Omics can be broken down into:

- Genomics: sequences the entire genetic profile of an individual patient;
- Epigenomics: measures modifications to DNA.
- Proteomics: measures all the proteins in a biological matrix, like plasma;
- Metabolomics: measures metabolites, substrates, and intermediary products found in metabolism;
- Transcriptomics: measures the set of all RNA transcripts in a cell or tissue, and
- Lipidomics: measures all lipids in a biological matrix, like plasma.

In companion diagnostics, there is often a single measurement, i.e., whether the patient has a particular marker in a tumor or in the blood. With -omics, thousands, and maybe even millions, of data points are generated from each sample. Consider that DNA alone has 3 billion base pairs. Whole genome sequencing files can be upwards of 100 gigabytes in size. Exome sequencing, which focuses on the 1-2% of DNA that encodes mRNA transcripts for protein synthesis rather than the whole genome, requires file sizes of 5 to 7 gigabytes. In comparison, the size of this book in Microsoft Word is around 0.05 gigabytes. Datasets this large require high-performance computing to analyze them in a timely manner. Furthermore, advanced numerical and computational analysis methods are

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



needed for data manipulation and extraction. There are entire fields, such as computational biology and bioinformatics, that study how to analyze large, complex datasets.

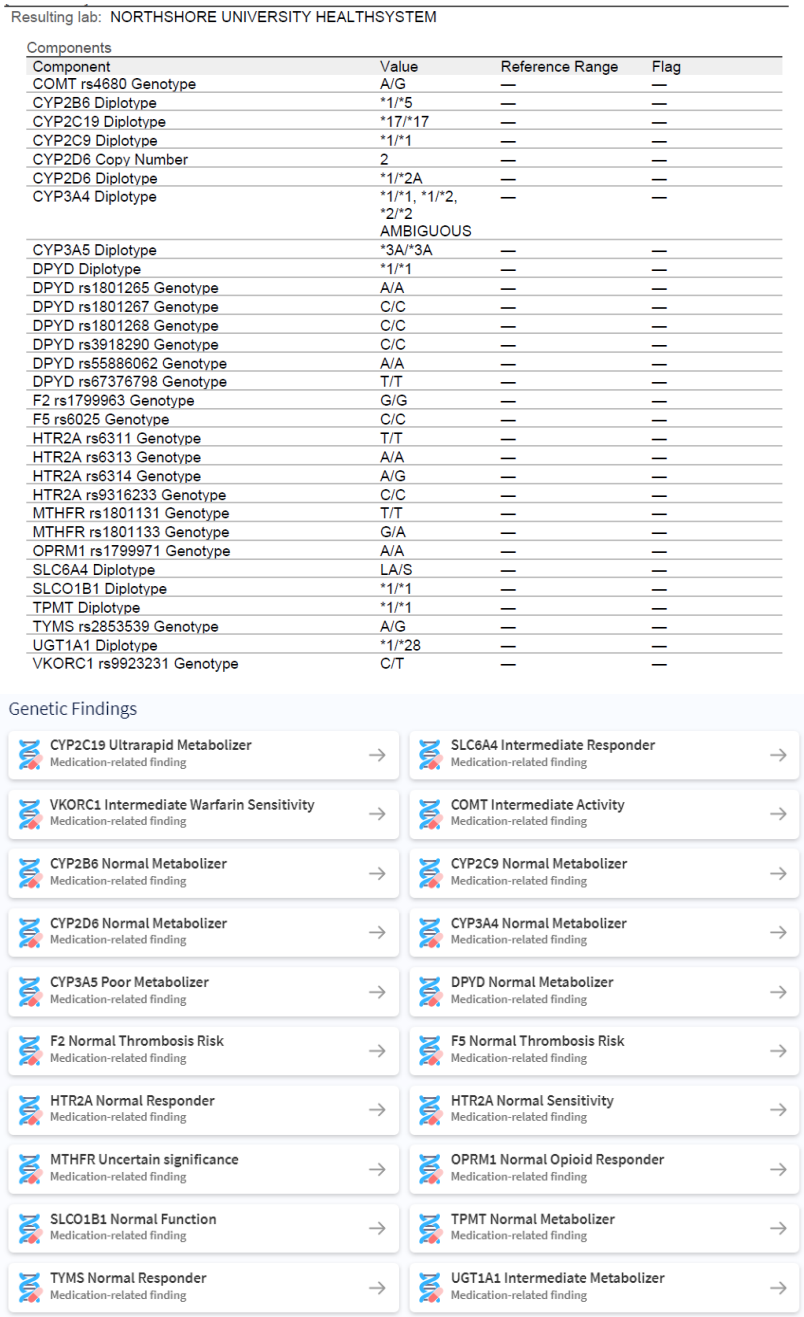
These methods, however, only provide part of the puzzle. To identify a subgroup of patients who respond better to a drug than others, you need to tie -omics data to outcomes. This requires linking such databases to outcomes databases, which may require access to hospital data systems and insurance databases. This creates huge issues regarding patient privacy and data linking across different databases. Companies, like Komodo Health, exist today to make these linkages available to researchers and other companies – for a price.

President Obama, in his 2015 State of the Union Address, proposed the Precision Medicine Initiative, a 1-million-person cohort of individuals who would donate their personal information, genetic information, and medical histories to science. The scope of this initiative would rival the Mission to the Moon that President Kennedy initiated in the 1960s. In 2018, the National Institutes of Health initiated the All of Us Program, which is expected to last at least 10 years. The goal is to link health and genetic information for 1 million individuals and make it available to researchers. Individuals who are typically underrepresented in scientific research will be actively recruited. As of 2020, ~350,000 individuals have donated their information to the program. As of yet, the long-term goals have not been reached. This database could potentially influence the scope of science for decades to come.

Although not being done anywhere near the scope it is being sold at, precision medicine is being done in clinical practice today, albeit in less ambitious forms. Some physician groups or hospitals offer their patients whole genome sequencing or simpler pharmacogenetic screens. Sequencing the entire genome for a patient can cost a few hundred dollars today compared to the thousands it cost just a few years ago. Pharmacogenetic screens are often comparable in cost but are focused more on drug metabolism polymorphisms (Figure 94). It has been argued that better dose selection can be achieved with these screening panels, leading to improved treatment response and reduced incidence of adverse events [362]. However, direct evidence from prospective clinical trials supporting these assertions is hard to find.

In some cancer patients, tumors are profiled, with 200 to 500 genes usually screened. Many hospitals now offer "tumor boards," which bring together oncologists, pathologists, bioinformaticians, ethicists, surgeons, pharmacists, and geneticists to identify the best treatment for a patient. Rather than treatment decisions being made by one or two physicians, tumor boards offer a range of opinions that may be more beneficial to patients. Cancer is largely organ-driven in its treatment, i.e., breast cancer, gastric cancer, and colon cancer. Still, molecular profiling of the tumor can help identify the alteration without focusing on the organ itself; for example, the tumor may harbor a BRAF mutation or an ALK gene fusion.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Resulting lab: NORTHSHORE UNIVERSITY HEALTHSYSTEM

Components

Component	Value	Reference Range	Flag
COMT rs4680 Genotype	A/G	—	—
CYP2B6 Diplotype	*1/*5	—	—
CYP2C19 Diplotype	*17/*17	—	—
CYP2C9 Diplotype	*1/*1	—	—
CYP2D6 Copy Number	2	—	—
CYP2D6 Diplotype	*1/*2A	—	—
CYP3A4 Diplotype	*1/*1, *1/*2, *2/*2 AMBIGUOUS	—	—
CYP3A5 Diplotype	*3A/*3A	—	—
DPYD Diplotype	*1/*1	—	—
DPYD rs1801265 Genotype	A/A	—	—
DPYD rs1801267 Genotype	C/C	—	—
DPYD rs1801268 Genotype	C/C	—	—
DPYD rs3918290 Genotype	C/C	—	—
DPYD rs55886062 Genotype	A/A	—	—
DPYD rs67376798 Genotype	T/T	—	—
F2 rs1799963 Genotype	G/G	—	—
F5 rs6025 Genotype	C/C	—	—
HTR2A rs6311 Genotype	T/T	—	—
HTR2A rs6313 Genotype	A/A	—	—
HTR2A rs6314 Genotype	A/G	—	—
HTR2A rs9316233 Genotype	C/C	—	—
MTHFR rs1801131 Genotype	T/T	—	—
MTHFR rs1801133 Genotype	G/A	—	—
OPRM1 rs1799971 Genotype	A/A	—	—
SLC6A4 Diplotype	LA/S	—	—
SLCO1B1 Diplotype	*1/*1	—	—
TPMT Diplotype	*1/*1	—	—
TYMS rs2853539 Genotype	A/G	—	—
UGT1A1 Diplotype	*1/*28	—	—
VKORC1 rs9923231 Genotype	C/T	—	—

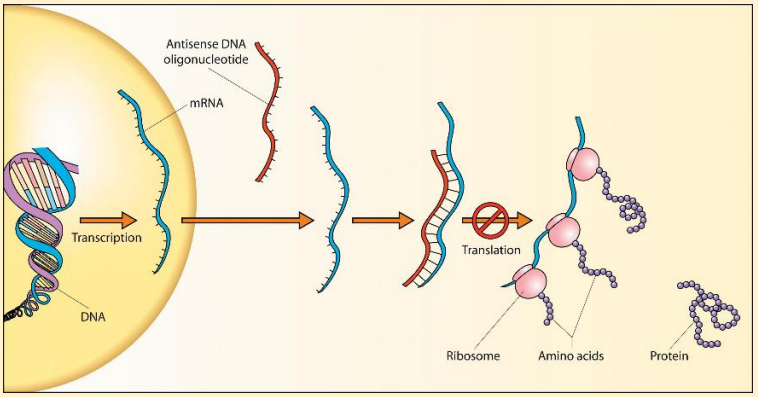
Genetic Findings

- CYP2C19 Ultrarapid Metabolizer — Medication-related finding
- SLC6A4 Intermediate Responder — Medication-related finding
- VKORC1 Intermediate Warfarin Sensitivity — Medication-related finding
- COMT Intermediate Activity — Medication-related finding
- CYP2B6 Normal Metabolizer — Medication-related finding
- CYP2C9 Normal Metabolizer — Medication-related finding
- CYP2D6 Normal Metabolizer — Medication-related finding
- CYP3A4 Normal Metabolizer — Medication-related finding
- CYP3A5 Poor Metabolizer — Medication-related finding
- DPYD Normal Metabolizer — Medication-related finding
- F2 Normal Thrombosis Risk — Medication-related finding
- F5 Normal Thrombosis Risk — Medication-related finding
- HTR2A Normal Responder — Medication-related finding
- HTR2A Normal Sensitivity — Medication-related finding
- MTHFR Uncertain significance — Medication-related finding
- OPRM1 Normal Opioid Responder — Medication-related finding
- SLCO1B1 Normal Function — Medication-related finding
- TPMT Normal Metabolizer — Medication-related finding
- TYMS Normal Responder — Medication-related finding
- UGT1A1 Intermediate Metabolizer — Medication-related finding

Figure 94. Sample report from the MedClueRx pharmacogenomics panel offered by Endeavor Health in Chicago. *The top* figure is the genotype results. *The bottom* figure is the phenotype results. Using this panel, 98% of patients have been found to have at least one polymorphism that may affect their patient care. This patient is normal in all regards, but they are an ultrarapid CYP2C19 metabolizer, which could affect their ability to metabolize drugs like omeprazole and sertraline. They are also have a VKORC1 C/T genotype, *which means they* have intermediate sensitivity to warfarin and are at increased risk for adverse events.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Knowing the tumor alteration may point towards drugs that work better on those alterations [363]. But, again, direct evidence from prospective clinical trials supporting this assertion is generally lacking, although literature reviews and meta-analyses suggest this is the case [364, 365].



Antisense oligonucleotides are short, synthetic, single-stranded RNA or DNA sequences that can bind to RNA sequences and alter protein expression [366].

Figure reprinted from Robinson R, CC BY 2.5 <https://creativecommons.org/licenses/by/2.5>, via Wikimedia Commons.

To date, precision medicine has focused on identifying the best drug for a patient, but researchers in Boston, across different institutions, took a different approach. Working together, they identified the genetic defect in a 6-year-old girl who presented with ataxia, blindness, seizures, and was developmentally regressed [366]. They discovered she had a form of a rare disease called infantile Batten's syndrome, which affects the brain and nervous system and is present in about 2 to 3 patients per 100,000. Specifically, the patient had a dysfunctional variant of the MFSD8 gene, which encodes a lysosomal membrane protein that transports various substrates. Lysosomes are the cell's garbage disposals, and without this integral protein, toxic waste products accumulate slowly, poisoning the patient.

The patient, who started to develop symptoms around 3 years of age, was not expected to live past age 10. In traditional precision medicine practice, the patient would be given a drug to treat Batten's disease. However, there are no approved drugs to treat Batten's disease. Instead, the researchers made a drug – just for her!

Building off an already approved drug for spinal muscle atrophy, the researchers thought that using the same approach might work for the patient, whose name was Mila. The researchers designed a 22-length antisense oligonucleotide (ASO) that they believed would restore normal function in the patient. They named the drug 'milasen' after the patient's name. In normal protein production, DNA is transcribed into RNA, which is then translated into a precursor protein; inactive regions of the protein are then removed (protein splicing), creating an active protein. Preclinical studies indicated that milasen more than tripled the normal level of MFSD8 protein splicing. Toxicology studies showed milasen to be generally safe at doses up to 42-times the dose to be administered in humans, except for some hindlimb weakness in some animals. Based on their preclinical package, the researchers filed an Expanded Access Investigative New Drug Application, which the FDA approved.

Milasen was administered directly into the spinal cord, with doses increasing every two weeks for 12 weeks. After dose escalation, two additional doses were administered, followed by repeat doses every three months. The total follow-up was about a year, during which no adverse events were observed. Milasen concentrations increased in cerebrospinal fluid in the brain, as measured at every visit for up to 200 days, and then plateaued thereafter. At baseline, the patient was having 15 to 30 seizures a day, each lasting 1 to 2 minutes. Over the course of the year, seizures decreased to 0 to 20 seizures a day, lasting less than a minute. While showing promising results, unfortunately, the disease was so advanced that Mila died at 10 years of age. Nevertheless, this groundbreaking work proved that a new drug could be tailored to a patient. Instead of trying to fit an octagon wood block into a round

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Are Redheads a Special Population?

There is evidence that redheads metabolize certain drugs differently and may require different doses than others [368]. Redheads may also exhibit different pain sensitivities compared to others and may respond more strongly to opioids than others [369]. They may also require a higher dose of local anesthetics, like lidocaine, for pain relief, possibly due to mutations in the MCR1 gene [370]. Does this mean that redheads are a special population?

hole, the researchers made a round block to fit the hole exactly.

And milasen was not a one-time thing. In 2023, the same researchers who developed milasen reported on a second ASO they developed for a patient with ataxia-telangiectasia (A-T), a rare childhood disease affecting 1 in 40,000 to 100,000 children and leading to movement and speech problems. A-T is a progressive disease that leads to death at around 25 years of age. A-T is caused by the loss of the ATM gene, which is involved in DNA damage repair. After meeting the founder of the A-T Children's Project, a nonprofit advocacy group for families with A-T children dedicated to finding a cure, Timothy Yu, the lead researcher for milasen, directed their research efforts toward an ASO for A-T. They reported results in 2023 after dosing a 2-year-old with A-T [367]. Although no clinical outcomes were reported, the drug, called atipeksen, appeared well-tolerated in the patient.

These successes led to the creation of several nonprofit organizations, such as the N=1 Collaboration (www.n1collaborative.org), dedicated to developing medicines for individuals. The Boston researchers never discussed how much it cost to make mirasen or atipeksen, but it must have been in the multimillion-dollar range. This raises many ethical and economic questions. Who will pay for this? The costs are not sustainable to do this for every patient. If it is paid for by the parents or some GoFundMe fundraiser, what about children who don't have rich parents or whose story is not as sympathetic as that of some other child's? Who then will pay for their drug? Does the government have an obligation to contribute?

While this may not seem like an economical way to develop drugs, that's how disruptive technologies work. At first, they are economically unfeasible, creating their own markets, and may not be as useful as other products. Still, over time, they get better as costs go down, eventually outperforming and then taking over the market from the alternatives. The classic example of a disruptive technology is the personal computer. When first introduced, it was unclear who would want one. They performed poorly, were expensive, and you couldn't save anything on them. But over time, they got cheaper, chips got better, and they eventually overtook mainframe computers to the point that almost every home now has a PC. It's not hard to envision a future in which personalized antisense oligonucleotide technology becomes readily available to patients with rare diseases.

Summary

There is a disconnect between the lay press and reality about personalized medicine, with the press tending to overhype personalized medicine, perhaps too early. It's easy to see why you would want personalized medicine. Who wouldn't want a drug that is tailored to the individual instead of a one-size-fits-all model that drug companies promote? To date, though, this has not panned out. The Human Genome Project has been criticized as a failure, or at least a failure to meet the hype, and, despite the entire human genome sequencing costing less than $1000 today, sequencing of patient DNA has not entered clinical practice, largely because it's not known what to do with all that data. The same holds for personalized medicine, which is more aspirational than inspirational.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The Perils of Plavix, or the Financial Consequences of Excluding Pharmacogenetic Information From the Label

Plavix (clopidogrel) is an antiplatelet medication used to prevent heart attacks and strokes in high-risk patients. Clopidogrel is a prodrug; it is metabolized to its active form by CYP2C19, which is a polymorphic enzyme. It is the active metabolite that binds to platelets, preventing platelet aggregation. Different ethnic groups have different proportions of polymorphisms. The *1 haplotype is associated with normal enzymatic function, whereas *2 and *3, which have reduced enzymatic function, are common among Asians and native Hawaiians and can lead to reduced clopidogrel efficacy. While not in the original label, in 2010, a black box warning was issued with regard to CYP2C19 polymorphisms:

WARNING: DIMINISHED ANTIPLATELET EFFECT IN PATIENTS WITH TWO LOSS-OF-FUNCTION ALLELES OF THE CYP2C19 GENE

See full prescribing information for complete boxed warning.

- Effectiveness of Plavix depends on conversion to an active metabolite by the cytochrome P450 (CYP) system, principally CYP2C19. (5.1, 12.3)
- Tests are available to identify patients who are CYP2C19 poor metabolizers. (12.5)
- Consider use of another platelet $P2Y_{12}$ inhibitor in patients identified as CYP2C19 poor metabolizers. (5.1)

In 2014, Holly Shikada, Attorney General for Hawaii and representing the State of Hawaii, sued Bristol Myers Squibb (BMS) and Sanofi, the manufacturers of Plavix, under Hawaii's Unfair or Deceptive Practices Law in regard to Plavix's label [371]. Since Hawaii has a higher proportion of CYP2C19 poor metabolizers, Shikada claimed that, at the time of approval, BMS/Sanofi were aware of the polymorphisms and that these patients would be less likely to respond to the drug. Shikada claimed BMS/Sanofi failed to highlight this in its label and warn prescribing physicians about this issue. Shikada also claimed BMS/Sanofi suppressed further research into the poor metabolizer issue and, in effect, misled the public. The state was seeking penalties of $10,000 per prescription written in Hawaii since 1998!

After the trial, the judge issued a ruling agreeing with the State that BMS/Sanofi engaged in deceptive and unfair acts or practices and imposing a $834 million penalty against BMS/Sanofi. Upon appeal to the Hawaii Supreme Court, the higher court ruled that BMS/Sanofi was unable to present a complete defense at trial, but upheld the deceptive practices ruling. The Hawaii Supreme Court vacated the $834 million judgment and remanded the case back to trial. Finally, in 2025, the two sides settled on a ~$700 million fine against BMS/Sanofi, with BMS/Sanofi not admitting any wrongdoing [372] [373].

Following the easy money, the Shikada outcome has led to other states following suit: Texas brought a similar lawsuit against BMS/Sanofi in 2025 [374]. In a broader sense, the outcome of this trial will undoubtedly lead to more consumer lawsuits against companies regarding efficacy and safety across different ethnic subgroups or demographics, raise questions about the ethical obligations for studying drugs across different subgroups, and raise questions about the disclosures required on drug product labeling.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Today, personalized medicine takes a simpler form that focuses on a single mutation or biomarker in deciding whether a patient is a candidate for the drug, i.e., does the patient have a particular mutation or some level of biomarker? These single tests are called companion diagnostics and are often developed in parallel with new drugs during the development cycle. Whereas companies were loath to use diagnostics for fear of reducing their potential market, companion diagnostics are now seen as a positive, net benefit for the company because the approval rate for drugs with companion diagnostics is higher than for drugs without a diagnostic, and are better for the patient, making it more likely they will be prescribed the drug should the diagnostic demonstrate they are a candidate for the drug. Companion diagnostics are expected to increase in use and be applied to more therapeutic areas, rather than being used mostly in oncology, as is the case today. The companion diagnostic market is expected to grow from 9.4 billion in 2024 to 22.4 billion in 2032.

How long it will take for companion diagnostics to morph into the type of personalized medicine hyped by the press remains to be seen. Harold Varmus, the former director of the National Institutes of Health, has said that genomic medicine will be a gradual, slow process that will take decades to achieve. Precision medicine will probably take just as long.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Appendix

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Appendix 1:
More on Estimating Human Doses from Animals and First-in-Man Dose Selection

One of the most difficult challenges in drug development is translating results from animals to humans, particularly converting drug dosages used in animal experiments to human doses, or vice versa. The first-in-man study, where humans receive the drug for the very first time, is one point in the development process where this translation occurs. It's important to pick a safe dose. Getting the wrong dose can lead to tragic consequences, as demonstrated by the TeGenero disaster in 2006. It is generally recognized that, in most cases, particularly with small molecules, simply translating the animal dose expressed as a weight (e.g., mg/kg) into humans will result in overdosing, although there are exceptions for certain drug classes, such as large proteins like monoclonal antibodies and gene therapy products. This appendix will focus on methods for translating animal doses to humans and selecting a starting dose for first-in-human studies.

Even translating mg/kg across animal species, such as using a mg/kg dose intended for a rat in a dog, can lead to tragic results. One example of cross-species tragedy is the story of Tusko, a 14-year-old male Asiatic elephant, which happened in 1962 [375]. Male elephants exhibit a behavior called 'musth,' a form of elephant madness that begins in adulthood (between 12 and 20 years) and can occur once or twice a year until the animal reaches 45 to 50 years old. There are no animal models for this condition. When 'musth' occurs, it can be quite dangerous for other elephants and their zoo handlers, so having an animal model to study the condition could be quite valuable. "Scientists[11]" at a zoo in Oklahoma got the idea that maybe LSD, a drug that induces psychosis in humans, could be used to mimic musth in elephants. Tusko was dosed with 0.1 mg/kg LSD or almost 300 mg of LSD based on his weight. The dose was chosen based on a 0.15 mg/kg dose used to induce a rage reaction in cats. In humans, just 100 to 200 μg is needed for effect. Tusko received about 2000-times the dose in humans. Within a few minutes of dosing, Tusko fell to the ground and died an hour later. What went wrong? West et al. failed to realize that although doses scale by weight within a species, they may not scale appropriately across species. Weight-based dosing frequently results in overdosing. Later analyses, using more appropriate dose translation methods, showed that a dose of 3-56 mg would have been more appropriate for Tusko [376].

It is now generally accepted that it is safer to scale doses across species by body surface area (BSA) rather than by weight. This dose is called the Human Equivalent Dose (HED) [377]. The core formula is:

$$\underset{\text{in mg/kg}}{\text{HED}} = \underset{\text{in mg/kg}}{\text{Dose in animals}} \left[\frac{\text{animal weight in kg}}{\text{animal BSA in m}^2}\right]\left[\frac{\text{human BSA in m}^2}{\text{human weight in kg}}\right]. \quad (39)$$

Everything inside the brackets in Eq. (39) is a species-dependent constant. For example, a 60 kg human has a BSA of 1.63 m^2. The equation can then be simplified to:

$$\underset{\text{in mg/kg}}{\text{HED}} = \underset{\text{in mg/kg}}{\text{Dose in animals}} \times \text{constant}. \quad (40)$$

[11] I use the term "scientists" in parentheses because one of the authors, Louis West, would later go on and become a psychiatrist of some notoriety. He would later work with the CIA on their LSD mind-control program.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Appendix 1 Table 1 Conversion of Animal Doses to Human Equivalent Doses Based on Body Surface Area			
Species	**Multiplication factor to convert animal dose in mg/kg to human dose in mg/m²**	**Multiplication factor to convert animal dose in mg/kg to human dose in mg/kg (constant in Eq. (40))**	**Multiplication factor to convert animal dose in mg/kg to human dose in mg (Eq. (41) using 60 kg weight)**
Mouse	3	0.081	4.9
Rat	6	0.162	9.7
Monkeys (cynomolgus, rhesus, or stumptail)	12	0.324	19.4
Dog	20	0.541	32.5
Notes: Calculations are based on a 60 kg human adult. HED = animal dose in mg/kg × (human weight in kg/animal weight in kg)$^{0.67}$			

To convert to a flat dose, Eq. (40) is simply multiplied by human weight:

$$\text{HED in mg} = \frac{\text{Dose in animals}}{\text{in mg/kg}} \times \text{constant} \times \text{human weight in kg}. \qquad (41)$$

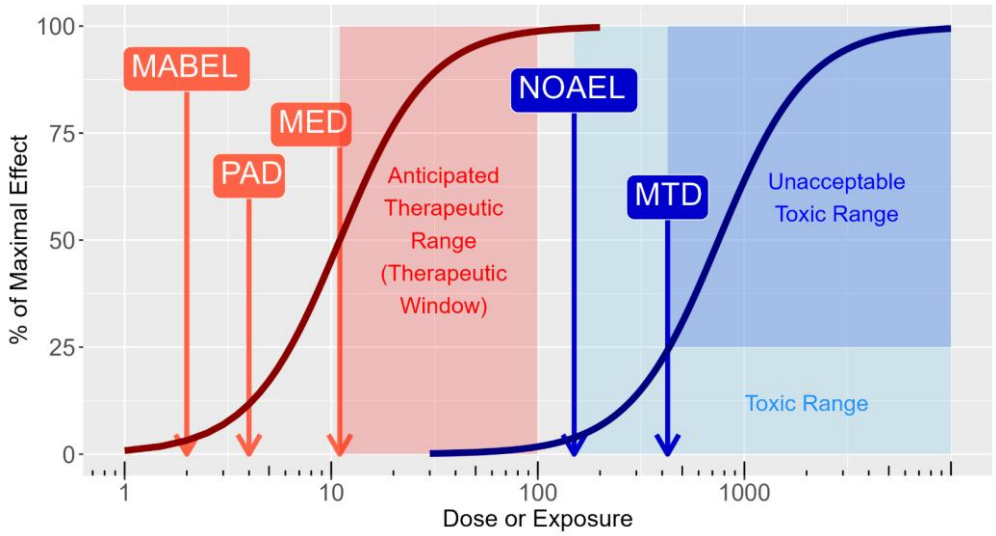
The constant[12] is species-dependent, e.g., mouse-to-human or rat-to-human. Appendix 1 Table 1 shows the conversion factors for common animal species. Suppose the pharmacologically active dose in mice was 30 mg/kg. The HED would be 147 mg (30 × 4.9), not 1800 mg (30×60) as would be estimated if the conversion was straight mg/kg in humans.

In first-in-man studies, the starting dose is based on the No Observable Adverse Effect Level (NOAEL) in toxicology studies. The NOAEL is the highest dose at which no statistically significant or biologically significant adverse effects are observed compared to control animals. The NOAEL can be contrasted to the Maximum Tolerated Dose (MTD), which is the highest dose that can be administered without *unacceptable* or dose-limiting toxicity. The MTD is by definition higher than the NOAEL. Suppose the NOAEL in rats was 100 mg/kg; the HED would be 970 mg (100×9.7). In first-in-man studies, where the drug has never been tested in humans, a safety factor is usually applied to protect human participants. Typically, a safety factor of 10 is applied. If a HED of 970 mg is estimated, then 97 mg (or 100 mg rounded up) would be administered. After application of the safety factor, the HED is renamed the Maximum Recommended Safe Dose (MSRD). Larger safety factors may be used depending on the drug's pharmacology. For example, immune stimulants may use a safety factor of 100 or more. These concepts, and others that are in this section, are illustrated in Appendix 1 Figure 1.

The starting dose in a first-in-man oncology study for a small molecule is slightly different from what was just described [378]. Because doses in oncology are typically pushed as high as possible while still being safe, and that it is unethical to administer doses to cancer patients that are known to not be pharmacologically active, instead of the using the NOAEL to calculate the starting dose in cancer

[12] The FDA guidance proposes naming the constant Km. This is confusing because Km is also the same constant used to define the Michaelis constant in enzyme kinetics and pharmacodynamic models. For this reason, the constant will just be named 'constant.'

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Appendix 1 Figure 1: Cartoon illustrating the various dose terms for a hypothetical drug. The dark red curve is the dose-response for drug efficacy. The dark blue curve is the dose-response for an adverse event of the drug.

patients, 1/6th the Highest Non-Severely Toxic Dose (HNSTD, the highest dose in animals that is not life-threatening or cause irreversible harm) in non-rodents, or 1/10th the Severely Toxic Dose (STD10, the lowest dose that causes lethality or life-threatening toxicity), in rodents is used. The choice between HNSTD or STD10 depends on which animal species best reflects the drug's action in humans and which results in the smaller dose estimate.

Suppose that for a new small-molecule cancer drug, 4-week toxicology studies were available in rats and monkeys, and the STD10 in rats was 40 mg/kg and the HNSTD in monkeys was 3 mg/kg. HEDs are calculated as 6.4 mg/kg for STD10 (40×0.16) and 0.96 mg/kg for HNSTD (3×0.324). For a 60 kg person, 1/10th rat STD10 is 38 mg, and 1/6th of the HNSTD is 9.6 mg. Of the two, the HNSTD is smaller, and a reasonable starting dose in human clinical trials would be 10 mg per week.

The MSRD approach makes certain assumptions, such as that the effect in animals at a given concentration will produce the same effect in humans at the same concentration, i.e., that animals and humans have the same sensitivity to the drug. This assumption implies that both animals and humans have the same degree of protein binding. If animals and humans do not have the same degree of protein binding, a pharmacokinetic approach may be needed. For example, if animals have 98% plasma protein binding and humans have 99%, the HED may need to be based on unbound AUC rather than animal dose. For oral drugs, it also implicitly assumes that bioavailability in animals is the same as in humans. If not, this may require a pharmacokinetic approach as well.

This MSRD approach is not appropriate for all conditions. If the drug is administered locally, such as by topical, intranasal, subcutaneous, or intramuscular routes, the HED should be normalized to the concentration, e.g., mg/area of application or the amount of drug at the application site. Further, proteins with molecular weights exceeding 100,000 Daltons should not be scaled by BSA; i.e., mg/kg in mice is the same as mg/kg in humans.

MSRD protects against toxicology-driven risk

PAD balances mechanism-based pharmacology with practicality

MABEL protects against unknown or exaggerated biology

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

The pharmacokinetic approach is another scaling approach; this time, the HED is based on a target exposure and is adjusted for species differences. Suppose the target AUC in humans is 100 ug*h/mL and the mouse clearance is 180 mL/h. Since AUC = D/CL and CL is proportional to weight with a 0.7 exponent, then

$$D = AUC \times CL_{animal} \left[\frac{\text{Weight in humans in kg}}{\text{Weight in animals in kg}} \right]^{0.7} \left[\frac{\text{unbound fraction in animals}}{\text{unbound fraction in humans}} \right]. \tag{42}$$

In this case, if the unbound fractions are 98% in mice and 99% in humans, then

$$D = 100 \text{ ug*h/mL} \times 180 \text{ mL/h} \left[\frac{60 \text{ kg}}{0.025 \text{ kg}} \right]^{0.7} \left[\frac{0.02}{0.01} \right] = 371 \text{ mg}. \tag{43}$$

A similar equation can be derived for the average steady-state concentration (Css,avg) equation by multiplying Eq. (42) by λ, the dosing interval:

$$D = C_{ss,avg} \tau \times CL_{animal} \left[\frac{\text{Weight in humans in kg}}{\text{Weight in animals in kg}} \right]^{0.7} \left[\frac{\text{unbound fraction in animals}}{\text{unbound fraction in humans}} \right]. \tag{44}$$

If the MSRD is estimated, it may be useful to also estimate the Pharmacologically Active Dose (PAD) in humans. The PAD is defined as the lowest dose that produces a pharmacologically meaningful and measurable response in humans, consistent with its mechanism of action. The PAD is not necessarily a clinically efficacious dose; the lowest clinically relevant dose is the Minimum Effective Dose (MED). There are situations where the NOAEL-based MSRD approach is inappropriate, and the PAD is more appropriate:

- PAD should be used if safety is tied to exaggerated pharmacology, i.e., doing what the drug is meant to do, rather than to off-target effects. Examples include immune agonists or modulators, cytokines, or highly potent enzyme inhibitors.
- PAD should be used when the human target biology is not well represented in animals, such as when a receptor is present in humans but not in animals, or when signaling pathways are altered in animals compared to humans. This is one of the proposed reasons for where TeGenero 1412 went wrong. Nonhuman primates lack CD28 on memory T-cells, and toxicology studies gave an overly optimistic safe dose.
- PAD should be used when the dose-response in animals is steep and small dose errors may lead to large biological effects.
- Drugs that exhibit nonlinear pharmacokinetics, where small changes in dose may lead to large changes in drug concentrations. Certain monoclonal antibodies are an example.
- If this is the first drug for a new class of drugs or when there is high uncertainty in a drug's mechanism of action, the PAD should be used instead of the MSRD approach.

Like the MSRD, a safety factor is often applied to the PAD, maybe a 1/3 or 1/10 reduction. The PAD is lower than the MSRD; it may be safer to start with the PAD than the HED.

For example, RO7119929, developed by Roche as an anti-cancer agent, is a small-molecule prodrug agonist of TLR7, which triggers the production of pro-inflammatory cytokines, such as interferon-α and interleukin-6, both of which are needed for an effective immune response [379]. Roche incubated RO7119929 in human white blood cells (mononuclear cells) to determine the minimum concentration required to increase cytokine secretion. They then measured unbound drug concentrations in monkey plasma and scaled the pharmacokinetics to human values using PBPK modeling. From this, they determined that a dose of 3 mg would be the PAD. The dose actually used in the first cohort of the first-in-man study was 1 mg. Results from that study demonstrated that the actual PAD was around 2 mg [380].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Related to the PAD is the Minimum Anticipated Biological Effect Level (MABEL), the lowest dose that produces a minimal, measurable biological effect in humans. The MABEL is not necessarily a safe dose by itself, nor does it represent a pharmacologically meaningful or clinically efficacious dose. There is no universal approach to determining a first-in-human dose using the MABEL, although it often accounts for human potency, such as using IC10 or EC10 values, receptor occupancy, and target expression levels in humans. BY definition, MABEL is smaller than the PAD and is the most conservative estimate of where biological activity may begin.

Factors to Consider in Calculating the MABEL

- Mode of Action
- Target Pharmacology
- Relevance of animal models
- Patient Population and level of acceptable risk
- Predicted human pharmacokinetics

Yadav et al. [381] reported on a recombinant targeted cytokine that combined an anti-programmed cell death 1 receptor antibody with engineered IL-15. For their first-in-man study, they examined the totality of data from toxicology studies, in vivo pharmacology, and animal pharmacokinetic studies, and compared it with a PAD derived from mouse studies. Using the most sensitive in vitro assays using blood mononuclear cells, they determined the EC50 for maximal effect. Since concentration is amount/volume, they multiplied the EC50 by the estimated human volume of distribution and arrived at a MABEL dose of 3 mcg/kg IV (EC50 was 0.070 mg/mL and the human volume of distribution was 43 mL/kg). The first cohort of the first-in-man study used a dose of 3 μg/kg.

The last method to estimate a starting dose is model-based, using systems pharmacology (SP) or other approaches. Systems pharmacology develops mathematical models of the biology of the system in question and the integration of drug effects in the model. More advanced models incorporate physiologically based pharmacokinetic models (PBPK) to account for drug distribution throughout the body. An example of this approach is with blinatumomab, a bispecific antibody that stimulates the immune system. Bispecific antibodies are modified antibodies that have two ligand-binding domains. Blinatumomab binds CD3, which is found on mature T-cells, and CD-19, which binds to B-cells. Hence, blinatumomab act as a kind of molecular bridge bringing together a cancerous B-cell and a T-cell, thereby activating the immune system to destroy the cancerous B-cell. In vitro cytokine release experiments can overestimate clinical activity and result in starting doses 500-times below activity [382]. Optimal activity occurs when there is just the right stoichiometric amount of bispecific antibody that can form the trimer CD3-bispecific-CD19 [383]. Too low a concentration of bispecific antibody can result in few trimers, and, surprisingly, too much bispecific antibody can result in only dimers of CD3-bispecific and CD19-bispecific. This latter phenomenon is called the hook effect. If one were to plot tumor cell killing vs. bispecific concentration, an inverted U-shaped curve would result. Phipps et al. [382] used this approach for an unspecified Astra-Zeneca bispecific antibody and showed that the starting dose for a first-in-man trial was 10-times higher using the trimer SP approach than with the MABEL, reducing the number of patients exposed to ineffective approaches.

In summary, first-in-man dose prediction can be challenging and, when not done properly, can compromise participant safety. There is no single right method for choosing the correct starting dose, as it depends on many factors. Nevertheless, many clinical trials are started every year without issue.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Appendix 2:
Design and Interpretation of Clinical Trials

There are two types of clinical research: interventional and observational. Observational research follows participants over time and observes what happens naturally; no treatments are administered. For example, a study may take an insurance database, compare members who have taken a statin to treat high cholesterol with those who have not, and estimate the rate of rhabdomyolysis (where muscle tissue breaks down and releases harmful chemicals into the blood) in the two groups. Observational studies cannot assign a cause; at best, they identify correlations or associations between groups and outcomes.

An interventional study actively assigns participants to a treatment to determine the cause-and-effect relationship. For example, a group of participants with high cholesterol could be separated into two groups – one group is treated with a statin, whereas the other is not. The incidence of rhabdomyolysis between the groups is then compared. A difference between groups is highly likely to be due to the treatment effect. The history of drug development is largely based on interventional trials, but, as mentioned earlier in the book, real-world data, which can be observational, is increasingly used to compare with active treatments or to support label claims. In this appendix, we will go into the weeds, exploring the designs, analyses, and interpretations of interventional clinical trials used to support drug registration, with an emphasis on pivotal registration studies. Whereas the book, in general, has focused on the drug label (what it is, what information it contains, how to read it, etc.), the goal of this appendix is to provide the reader with the background to read and interpret publications of clinical trial results.

What do you call participants in a clinical trial? In reviewing results from clinical trials, different words are used to describe the people who enroll in the study. It used to be that 'volunteers' described healthy subjects who enrolled in a clinical trial, while 'patients' described those with an illness who enrolled in a clinical trial. That changed, and they were both referred to as 'subjects.' That has changed again, and now the most recent terminology is 'participant' [13], which is consistent with the ICH's E6(R3) Good Clinical Practice Guideline and the Declaration of Helsinki.

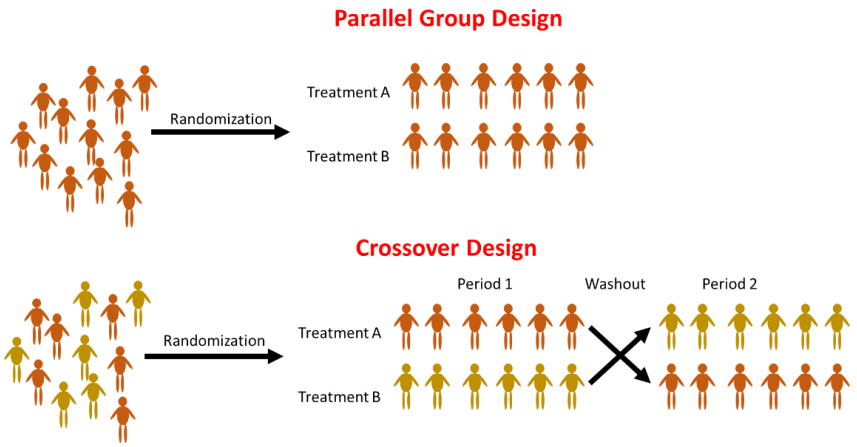
Phases of Drug Development Refresher

Phase 1 studies are designed to answer specific questions, such as: what is the effect of food on drug absorption, does Drug X interact with Drug Y, and what is the effect of Drug X on cardiac repolarization. There are generally two experimental designs used in Phase 1 studies: parallel group and crossover. In a parallel-group design, there are multiple treatment arms, and subjects are randomized to one of them. For example, in a food effect study, participants may be randomized to receive a single dose of the drug in either the fasted or fed state.

Phase 2 moves away from Phase 1, where "can we give this drug safely" is the primary concern, to one where we ask "does this drug work and at what dose?" Phase 2 is divided into Phase 2a, which demonstrates Proof of Concept that the drug works in patients, and Phase 2b, which identifies the

[13] I have probably not been consistent in this book with regard to terminology. It should be recognized that the terms subject, volunteer, and participant will be used interchangeably throughout the book. To keep track of healthy volunteers and patients, this book will use those terms instead of 'participants' whenever possible.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026



Appendix 2 Figure 1: Cartoon of a parallel group and crossover design for two treatments: A and B. Participants are color-coded in the crossover design to illustrate treatment assignment.

optimal dose at which the drug works in patients.

Phase 3 confirms that the drug works as intended in the study population. As such, these are referred to as pivotal, registrational, or confirmatory trials.

Phase 4 studies are conducted after the drug is approved for marketing by regulatory agencies. These studies may be additional Phase 1, Phase 2, or Phase 3 studies in the same population for which the drug was approved, or in new populations to gain additional indications.

Common Clinical Trial Study Designs

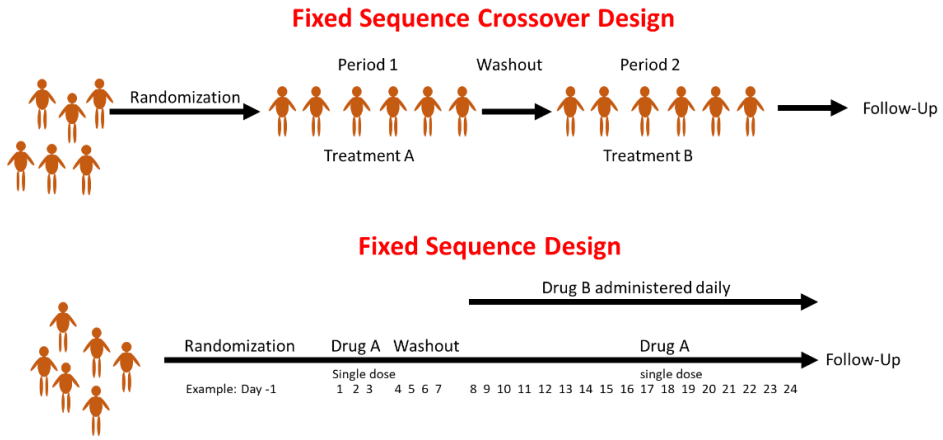
To be clear, there is no single standard clinical trial design for a given phase of drug development. There are many choices for Phases 2 and 3, fewer for Phase 1, but understanding these study designs will help interpret clinical trial results more effectively.

Most studies in drug development are superiority trials, in which the objective is to show that one treatment is better than another or a placebo.

Parallel Group Designs. The most commonly used study design is a randomized parallel-group design, in which participants are randomly assigned to one of two or more treatment groups. The groups are often one or more doses of an active treatment group and a comparison group, which is usually a placebo or standard of care. An example is presented in Appendix 2 Figure 1. In this example, there are two treatment groups, A and B, and participants are randomly assigned to receive either Treatment A or Treatment B. The biggest drawback of the parallel-group design is variability in the study endpoints across subjects. Large variability requires larger sample sizes to detect a difference between groups, i.e., there is low signal-to-noise. There is also an important assumption that randomization will minimize any bias between groups. For example, older participants may be more likely to experience an adverse event. If more elderly participants are assigned to one group over the other, the incidence of side effects may not be properly estimated. One advantage of its design is its simplicity: participants are followed for a fixed period, and then the study is completed. There are no modifications to the study design during the study.

A modification of the parallel group design is a stratified or blocked randomized parallel group design. Randomization is a key component of a parallel-group design, and it is assumed that, after randomization, participants are balanced with respect to factors that might influence treatment response. Stratification ensures balance across key prognostic variables, like disease severity, age, biomarker subgroup, and/or prior therapy. For example, suppose age was considered to influence

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Appendix 2 Figure 2: Cartoon of a fixed sequence crossover design and a fixed sequence design often used to study drug-drug interactions.

the outcome. You don't want to leave to chance that one treatment group has more elderly participants than the other; this may influence your estimate of the treatment effect. In a stratified block design, you may have two groups: 20 to 50 years and 50 to 80 years. Each participant is assigned to a stratum based on their age, and the randomization schedule for that stratum is used for that participant. When a stratum is full, no further participants are randomized into it. This approach ensures balance across age groups, reduces bias, and is more likely to detect treatment differences between groups.

Crossover Designs and Their Modifications. In a crossover design, each subject receives two or more treatments in a specific sequence during the study. In contrast, in a parallel-group design, each subject receives only a single treatment. For example, in a 2-period crossover, half of the subjects are randomized to receive one treatment in the first period, and the other half receive the other. This is followed by a washout period to allow the drugs to clear from the body and to avoid any carryover effect from Period 1 to Period 2. The experiment is then repeated in Period 2 after switching the treatments. For example, half the participants may receive Treatment A in Period 1 and then switch to Treatment B in Period 2. In the second group, half the participants receive Treatment B in Period 1 and then switch to Treatment A in Period 2. This is illustrated in Appendix 2 Figure 1.

Crossover designs are useful because each subject can serve as their own control, and the number of subjects needed to study is generally reduced compared to the parallel-group design. But, sometimes the crossover is impractical because, if there are many treatment groups, it may not be practical for every subject to receive every treatment; you cannot administer all the treatments to participants because the length of each period is sufficiently long that it is impractical for subjects to complete all the treatment periods. Or sometimes, if the washout is of insufficient length, Period 1 and Period 2 are not independent, and effects from Period 1 carry over into Period 2. Drugs with shorter, reversible effects are good choices for crossover designs.

Bioequivalence studies often use crossover designs in which each subject receives a different formulation of a drug product in separate periods, thereby assessing the pharmacokinetics of the formulations within the same subject, reducing variability, and increasing the probability of detecting a difference between the different formulations.

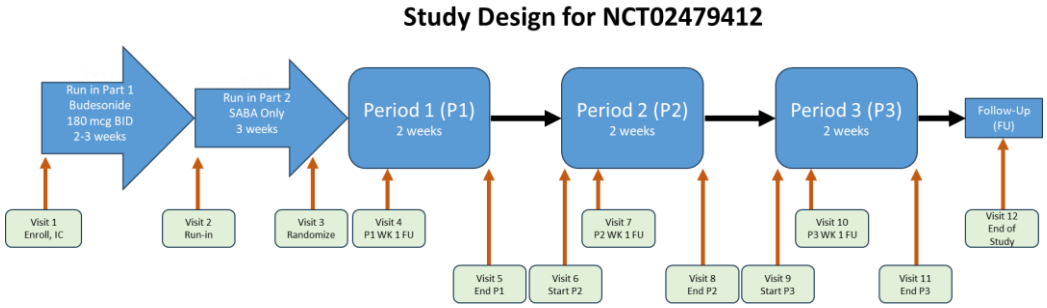
A modification of the crossover design is the fixed-sequence design, in which all subjects receive Treatment A, a washout, and then Treatment B in a fixed sequence over time. This design allows a within-subject comparison of the effect of Drug A vs. Drug B, but assumes there is no carryover of effect from Period 1 into Period 2. A modification of the fixed-sequence design can be seen in certain

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

drug interaction studies, in which Drug B is administered and washed out; Drug A is then started and administered repeatedly until steady-state is achieved, at which point Drug B is administered again in combination with Drug A. In this manner, the effect of Drug B alone and Drug B in combination with Drug A can be estimated. These designs are illustrated in Appendix 2 Figure 2.

In single-ascending-dose designs, which are common in first-in-man studies, a modified crossover design called a leapfrog design was popular years ago but is rarely seen today. Suppose a study has planned six treatments: 100, 200, 400, 600, 800, and 1000 mg. A leapfrog design would be one in which one group receives 100, 400, and 800 mg, and another group receives 200, 600, and 1000 mg.

An example of a Phase 2a crossover is the study of velsecorat in patients with asthma (NCT02479412) [384]. In this Phase 2a study, there were four treatments: placebo and three doses of velsecorat (58, 250, and 800 μg). Patients were studied over three periods, each lasting two weeks, with a three-week washout between periods. Before Period 1, all patients received budesonide for two-to-three weeks, followed by a short-acting beta-agonist (SABA) to reduce background inflammation variability and ensure everyone had similar levels of disease control before the start of the experimental treatments. At the beginning and end of each period, patients were assessed for their lung function. A schematic of the design is as follows:



This study used an incomplete block design. Because there were four treatments, a study with that many treatment periods and run-in periods would be too onerous and difficult to complete. Instead, patients are randomly assigned to one of nine treatment sequences; they do not receive all treatments; rather, they receive a subset. The exact treatment sequences used in this study were not reported, but an example block design would be:

Sequence	Treatment Received		
	Period 1	Period 2	Period 3
1	Placebo	58 mcg	250 mcg
2	Placebo	250 mcg	800 mcg
3	Placebo	800 mcg	58 mcg
4	58 mcg	Placebo	250 mcg
5	250 mcg	Placebo	800 mcg
6	800 mcg	Placebo	58 mcg
7	250 mcg	58 mcg	Placebo
8	800 mcg	250 mcg	Placebo
9	58 mcg	800 mcg	Placebo

Under this design, the treatments are said to be balanced, in that each active treatment is received six times, and the placebo is received in every treatment sequence; therefore, this is referred to as a balanced incomplete block design.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

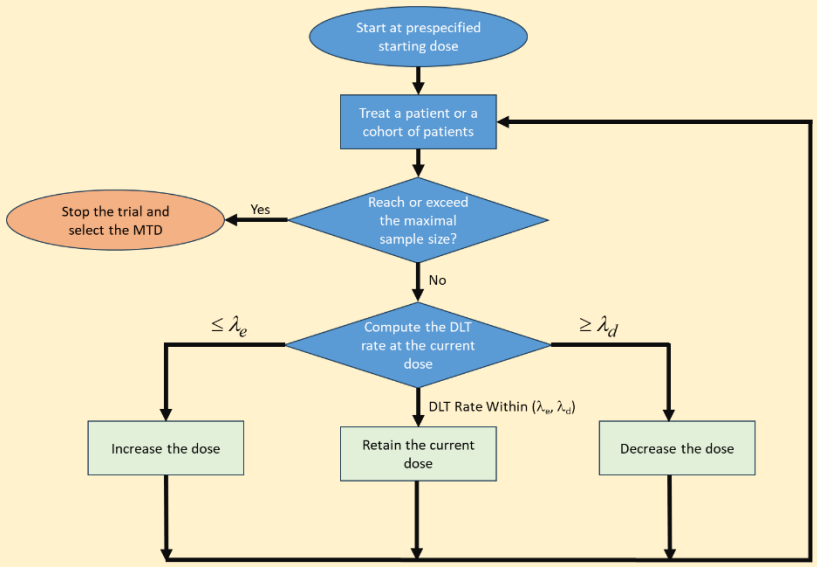
BOIN Designs in a Nutshell

In a BOIN design, typically three patients are enrolled at each dose level. If the total number of enrolled patients has not met the maximal sample size for the study, the Dose Limiting Toxicity (DLT) Rate is calculated as

$$\text{DLT Rate} = \frac{\text{Total number of patients who experienced a DLT at the current dose}}{\text{Total number of patients treated at the current dose}}.$$

If the observed DLT rate is less than the prespecified dose-escalation probability λ_e, the dose is increased. If the DLT rate exceeds the prespecified dose-deescalation probability ?, the dose is decreased. And if the DLT rate is within the window (λ_e, λ_d), the dose is maintained, and another patient is enrolled. The escalation and de-escalation probabilities are derived to minimize the probability of making an incorrect decision regarding increasing or decreasing the dose.

For an overall target toxicity rate of 25%, λ_e and λ_d are 0.197 and 0.298, respectively [385]. Hence, the BOIN design is transparent. If the observed DLT rate for a cohort exceeds 0.298, the next dose will be lower than the current dose, or if the current DLT rate is less than 0.197, the next cohort will have a higher dose. Computer simulations have shown that the BOIN design outperforms the 3+3 design at selecting the MTD, lowering the risk of overdosing patients, and that the 3+3 design is a special case of the BOIN design.



The BOIN design has been modified to account for different types of toxicities. A standard BOIN design requires that all patients complete the evaluation period before a decision is made to escalate, de-escalate, or maintain the dose. The TITE-BOIN (time-to-event BOIN) design is a BOIN modification that accounts for the time to DLT, allowing dose decisions to be made while waiting for late-onset toxicity to occur, as might occur with immunotherapies. Another design, the MT-BOIN, allows multiple toxicities, unlike the standard BOIN, which supports only a single toxicity. The generalized BOIN design allows for non-binary toxicity, such as using a toxicity score. Instead of a decision rule based solely on toxicity, the BOIN-ET design integrates safety and efficacy in dose-finding and aims to identify the optimal biological dose. There are many other types of BOIN designs.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

A special type of crossover design is a rescue design, in which patients are randomized to active treatment or to another treatment, which may be a placebo or another drug. If a placebo-treated participant fails to respond or meets predefined worsening criteria, the participant crosses over to active treatment. The crossover is conditional; participants only crossover if they get worse or fail treatment. An example of a rescue crossover is the EMPOWER-LUNG-1 study (NCT03088540) [386]. Patients with non-small cell lung cancer were randomized 1:1 to cemiplimab 350 mg intravenously every three weeks for two years or the investigator's choice of chemotherapy. Patients randomized to receive chemotherapy were allowed to crossover to cemiplimab if their cancer worsened while on chemotherapy (disease progression). After three years, about 75% of chemotherapy patients crossed over to cemiplimab. Rescue crossovers are not without problems, however. They can lead to problems with interpretation, particularly with time-to-event endpoints, and to imbalance across treatment groups.

Another type of crossover design is a randomized withdrawal design, which is similar to a rescue design but in the opposite direction. Patients all start on active treatment, and after a period of time, some or all are crossed over to placebo treatment to test maintenance of the active treatment effect. If placebo patients get worse, they are crossed back over to active treatment. Such a design is efficient in minimizing placebo treatment. It can be used as a type of enrichment design, focusing on responders or non-responders, but is hampered by high dropout and missing data. An example of this study design was reported by Tandon et al. on the effectiveness of lurasidone as a maintenance treatment for schizophrenia [387]. In that study, patients received 40 to 80 mg lurasidone once daily for 12 to 24 weeks. Patients who maintained clinical stability for 12 weeks or more were randomized to receive either placebo or 40-80 mg lurasidone for an additional 28 weeks of treatment. Because only responders cross over to the placebo arm, the study design enriches the treatment population by excluding non-responders before randomization.

Lastly, in some diseases, such as Alzheimer's disease, Parkinson's disease, rheumatoid arthritis, and chronic obstructive pulmonary disease, a type of crossover design called a delayed-start study is sometimes used [388]. In this study design, patients are randomized to active treatment or placebo, and after a period of time, the placebo group is switched to active treatment; this latter group is referred to as the delayed-start group. For these illnesses, it is important to distinguish between symptomatic treatment (which temporarily improves symptoms) and disease-modifying treatment (which slows or alters the course of the disease). If the early-starting group remains better than the delayed-start group even after both are on the drug, this suggests that the drug is disease-modifying rather than just providing symptomatic relief. This design is not often used and is more often seen in Phase 3 than in Phase 2.

Enrichment Designs. Enrichment designs have been discussed previously; at their core, they increase the probability of detecting a treatment effect by focusing on patients most likely to respond to the drug. This can be done through screening, such as only enrolling patients with a particular biomarker, or through a screening period in the clinical trial that weeds out non-responders, leaving only a sub-population of participants most likely to respond to the drug. By focusing on a subgroup of responders, variability among participants can be reduced, making clinical trials more efficient and cost-effective.

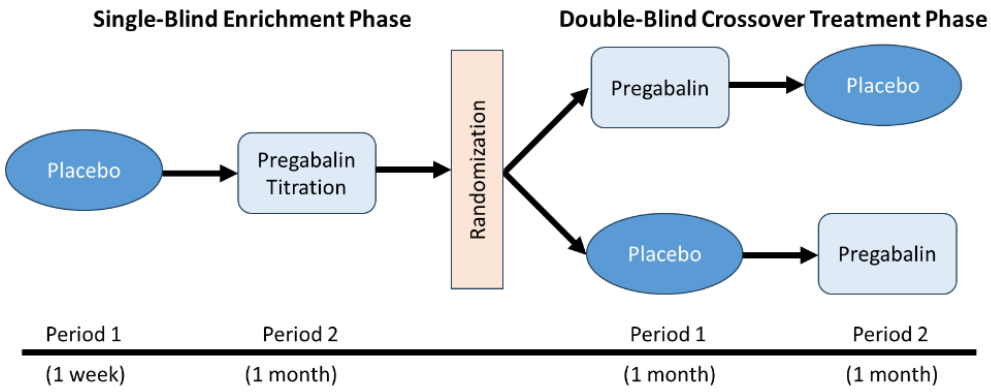
Enrichment designs generally fall into two types: prognostic enrichment and predictive enrichment. Prognostic enrichment selects patients who are more likely to experience the event of interest in event-driven studies or to have a substantial worsening of condition in continuous endpoint studies. Predictive

The difference between prognostic and predictive biomarkers can be confusing. Prognostic biomarkers predict disease outcome independent of treatment and can help identify patients at high or low risk. Examples of prognostic biomarkers are the stage or grade of tumor (Grade 3 colorectal, for example), or high CRP is predictive of major future cardiac events. A predictive biomarker indicates the likelihood of response to a specific treatment and identifies patients more likely to benefit from it; for example, HER2 overexpression predicts benefit from trastuzumab.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

enrichment is intended to increase the magnitude of the treatment effect. In both cases, prognostic or predictive indicators could include demographics, laboratory values, medical history, genomic or proteomic markers, or biomarkers. Such prognostic or predictive indicators are typically dichotomous, such as the presence or absence of a biomarker, or enrolling only patients with a systolic blood pressure of 160 mm Hg or higher for a new drug to treat hypertension. The entire chapter on Personalized Medicine was built around the idea of using companion diagnostics as predictive biomarkers.

A common predictive enrichment strategy, sometimes called response enrichment, is to enroll all patients who meet the inclusion/exclusion criteria and to include a placebo run-in period to weed out those with a strong placebo response. An example of this design is the study by Gonzalez-Duarte [389]. Patients with prediabetes, a milder form of glucose dysmetabolism, experience nerve pain (neuropathy) in less than 10% of cases. They sought to answer the question of whether prediabetic patients should be treated any differently than diabetic patients for their nerve pain. In this case, pregabalin is the drug of choice. To answer this question, they used an enrichment strategy with a single-blind enrichment phase followed by a double-blind active-treatment crossover phase. The enrichment phase was designed to reduce the placebo response rate by excluding patients who were likely to respond to a placebo. The design is as follows:



Using this design, they demonstrated that pregabalin was effective at treating prediabetic neuropathy.

A modification of this design is a randomized withdrawal design, which was described in the crossover section. Under this design, participants are assigned to active treatment during the prescreen period, and those who respond are randomized to receive either a placebo or to continue on active treatment. Such a design would be useful for long-term drug administration, where a placebo control arm would be unethical or impractical to conduct.

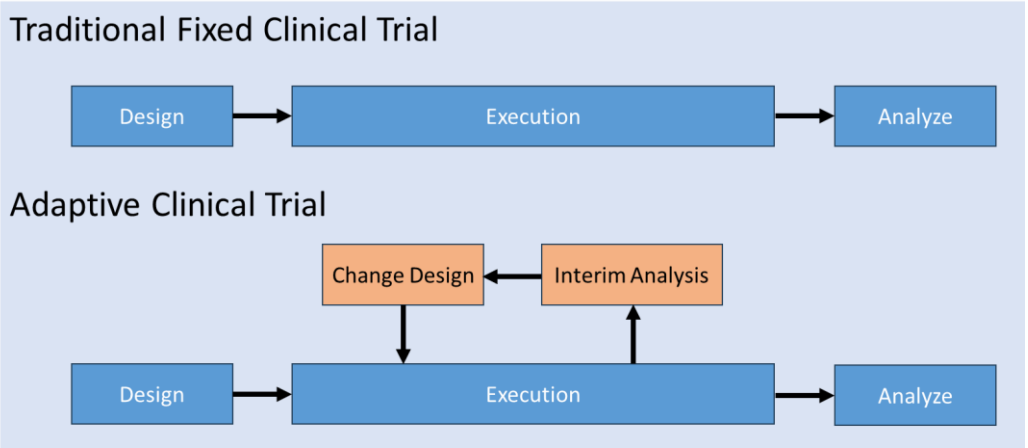
The FDA has issued guidance on this topic [390], and it contains many examples. The reader is referred there for further details on this study design.

Adaptive Designs. In a traditional randomized controlled clinical trial, participants are assigned to a treatment and followed until the end of the study. Treatments are fixed. Further, design elements such as sample size and dose are fixed. The trial is not altered during the study. Adaptive designs allow the design elements of a clinical trial to change as the study progresses, based on results from interim analyses conducted before the trial's completion.

For instance, suppose a clinical trial is done using a placebo and two doses of a drug. Halfway through a trial, it may be found that one particular dose level is ineffective; that dose level may then be terminated, and the trial continues with the remaining two treatment arms. Another example is sample size. In determining how many subjects are needed for a clinical trial, the variability of the outcome of interest in the population of interest must be known. Frequently, the sample size is calculated using an estimate of the variance. After observing the actual outcome of interest during

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

the clinical trial, it may be that the original estimate used to plan the trial was too small, and a larger sample size than originally planned is needed. Adaptive designs allow for these scenarios.



Adaptive designs, by their nature, are more flexible than their fixed design counterparts, leveraging data collected and knowledge gained during the trial's ongoing phase. However, this flexibility comes with a price, as care must be taken to avoid biasing the trial in favor of one particular treatment. Changes to the trial cannot be made willy-nilly after seeing the interim results. Adaptation plans with firm decision rules must be in place before the trial begins, and any interim analyses should be conducted by an independent group that is completely unblinded to all clinical trial data and has no financial or other stake in the trial's outcome. This may be a Data Safety Monitoring Board (DSMB) or an Interim Data Monitoring Committee (IDMC). Further, once a design change is made, special statistical methods must be used to account for it. Failure to account for any design change in the statistical analysis can lead to biased statistical results.

One type of adaptive design is a futility analysis, in which an interim analysis is conducted to determine whether it is worthwhile to continue the study. Given the interim data, the chances of seeing a statistically significant result at the end of the trial may be small. In this case, it is unethical to continue enrolling participants, and the trial is stopped for futility. Occasionally, the opposite is reported – where one treatment is so superior to the other that it is unethical to enroll participants in the other arm. Such a design, in which a decision to stop treatments is based on efficacy, safety, or futility, is also called a group sequential design.

Adaptive designs are not limited to Phases 2 and 3; they are also used in Phase 1. Recall that the first-in-man study is designed to find the highest safe dose that can be administered to humans. In a traditional design, the doses are fixed, e.g., volunteers will be administered 50, 100, 200, 300, 400, or 500 mg. If no toxicities are observed, 500 mg is considered the highest dose. In healthy volunteer studies, volunteers will be randomized to receive either the drug or a placebo; e.g., six volunteers will be randomized to receive the drug and two to receive the placebo at each dose level. In oncology, a common design is the 3+3 design, which is simple, algorithm-based, and easy to understand, enrolling three patients at each dose level. If no patient experiences toxicity, the dose is escalated to the next level. If one patient has toxicity, three more patients are enrolled at that dose. If none of those three new patients have toxicity, the dose is escalated. If more than two patients experience toxicity at a given dose, that dose is defined as the Maximum Tolerated Dose (MTD), and the dose is then reduced to identify the Recommended Phase 2 Dose (RP2D). Often, once an RP2D is identified, a dose-expansion cohort is used, enrolling 10 or more participants, to better confirm that this is indeed the optimal dose to take into Phase 2. This design in oncology is not very efficient at finding the MTD.

Newer experimental designs use adaptive designs to reach the MTD more efficiently by using

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

mathematical modeling. Designs like the continual reassessment method [391], Bayesian Optimal Interval Design (BOIN) [392], and backfill i3+3 designs [393] are all different types of adaptive designs that still use a 3+3 design but, instead of fixed dose levels, basically evaluate each dose level after they accrue and use mathematical models to predict the probability of adverse events at the next proposed dose level. If that predicted probability is too high, a lower dose is proposed. These designs have been described as more ethical than fixed-dose escalation studies because they may treat more participants at a potentially effective dose while protecting them from doses with a higher incidence of adverse events.

The use of adaptive designs is not without cost. As you increase the number of statistical tests performed in a study, the more likely you are to have a Type I error, where a hypothesis is falsely declared to be statistically significant. In a typical study, using an alpha (α) level of 0.05 for significance (i.e., the test is declared statistically significant if the p-value from the statistical test is less than 0.05), there is a 1 in 20 chance (5%) of falsely declaring a statistically significant result. An interim analysis yields two statistical tests: one at the interim and one at the end of the trial. Hence, the chance of a Type I error increases from 5% to 9.75%[14]. Hence, in adaptive designs, the error rate must be maintained at 0.05, even with multiple tests. This is done by splitting the α, across all the interim and final analyses. For example, an α-level of 0.0294 could be used for the interim analysis and an α-level of 0.0294 at the final analysis; this is the Pocock method. There are other ways to split the α, but whatever method is used, the α is protected and maintained at 0.05.

Much more can be said about adaptive designs; this is currently a hot topic in clinical trial design. The reader is referred to guidance issued by the FDA [394] and a primer on the topic [395].

Factorial Designs. Factorial designs are useful when the combined effect of two or more treatments is of interest. In a factorial design, participants are randomized to every possible combination of treatments. For instance, suppose there are two treatments, A and B, each having a low and a high dose. A factorial design would have four treatment arms:

1. Low A with low B,
2. High A with low B,
3. Low A with high B, and
4. High A with high B.

Under a factorial design, the effects of A and B can be estimated, as can the effect of A+B, which can then be tested to see if there is an interaction between A and B where the combination effect is greater or less than the sum of its parts. An example of a factorial design can be found with mirabegron and solifenacin, two drugs used to treat overactive bladder (OAB). It was thought that greater efficacy may be achieved by administering these two drugs in combination. Abrams et al. [396] tested this hypothesis using a factorial design in a Phase 2 study by examining in patients with OAB the mean volume of urine voided (MVV) and the number of micturitions in 24 hours for each treatment combination. A total of 1304 participants were administered either a placebo, monotherapy with solifenacin 2.5, 5, or 10 mg, monotherapy with mirabegron 25 or 50 mg, or combination therapy with solifenacin 2.5, 5, or 10 mg and mirabegron 25 or 50 mg, once daily, for 12 weeks. A total of 12 treatment groups were examined. The results showed that combination therapy was more effective than either treatment alone. Based on these results, the label for mirabegron (Myrbetriq) states that:

- "The recommended starting dose of MYRBETRIQ is 25 mg orally once daily, either alone or in combination with solifenacin succinate 5 mg orally once daily."

When done correctly, factorial designs can be analyzed using low-level response surfaces to identify maxima or minima, which may indicate synergy or interaction between treatments.

Noninferiority and Equivalence Trials. In a typical parallel-group design, we aim to show that

[14] The error rate making at least one Type I error for multiple tests is $1-(1-\alpha)^n$, where n is the number of statistical tests.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

our drug is superior to the comparator treatment (i.e., a superiority trial). In a noninferiority trial, we aim to show that a treatment is no worse than the comparator.

Noninferiority trials are more complicated than other designs, like a parallel group design, where the objective is to demonstrate a difference between groups, because it is impossible to prove that two treatments are identical in every respect. It is, however, possible to show that one treatment is not worse than the other treatment by an acceptably small margin with a given degree of confidence [397]. Whereas in a superiority trial, an intent-to-treat (ITT) population is preferred, dropouts, missed doses, and protocol violations may bias results toward noninferiority. For that reason, a per-protocol population may be preferred in the analysis of noninferiority trials.

It should also be pointed out that failure to reject the null hypothesis of no difference in a superiority trial does not mean that the treatment is noninferior. The adage, "The absence of evidence is not evidence of absence," applies here. The trial may have had too few participants to detect a difference if one were truly present, i.e., it was underpowered. The scientific literature is littered with underpowered clinical studies that have failed to detect a difference between two treatments. This does not mean that the treatments are the same.

In a noninferiority trial, the goal is to show that a treatment is no worse than a comparator. But in an equivalence trial, we want to show that a treatment is neither better nor worse than a comparator. We have already seen an example of an equivalence trial in bioequivalence testing, in which the generic drug aims to demonstrate that its rate and extent of absorption are the same as those of the brand-name product.

Statistically, for a noninferiority trial, we must show that the lower bound of the confidence interval for the difference between the treatment of interest and the comparator is greater than $-\Delta$, which is called the noninferiority margin. For an equivalence trial, the entire confidence interval for the difference must be contained within $\{\Delta_L, \Delta_U\}$[398]; the upper and lower margins are denoted U (upper) and L (lower), respectively, indicating that the bounds need not be symmetric. Choosing Δ is one of the most important decisions in a noninferiority or equivalence trial and usually requires discussion and agreement with regulators. It should be noted that different regulatory agencies may require different noninferiority margins. The choice of Δ must be defined prior to the start of the study and is not purely a statistical choice, but also requires clinical judgement.

Clinical Trial Elements

Design considerations for pivotal trials were introduced in the Phase 3 section of the chapter on Drug Development. This section will address elements that define a clinical trial.

Protocol. Under the Declaration of Helsinki, all clinical trials must have a protocol that details the objectives, all procedures for the study, and a semi-detailed description of how the data will be analyzed. All procedures must be described clearly and precisely. The procedures themselves should minimize risk and complexity to participants. Many published clinical trials include the protocol as a supplemental file.

Protocols must be approved by an independent Institutional Review Board (IRB) or Ethics Board. The FDA does not review all clinical trial protocols. The FDA typically reviews first-in-man studies, dose-escalation studies, Phase 2 and Phase 3 registrational studies, and major protocol amendments, such as new doses or populations. Also, the protocol must be sent to regulatory agencies in every country where a clinical trial is underway. For this reason, companies typically prepare a global master protocol and country-specific protocols to address concerns or laws in each country. And just because the FDA approves a protocol does not mean all other countries will as well.

Should it be necessary to make changes to the protocol during the study, a Protocol Amendment can be submitted to the IRB for approval. Sometimes, small changes to the protocol, such as clarifying procedures that do not affect participant welfare or safety, can be made via an

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

administrative letter. When in doubt, obtain IRB approval. Protocols with a large number of amendments (usually more than four) or amendments made late in the clinical trial raise red flags from regulators because they can indicate poor planning, potential bias towards favoring one treatment over another, operational bias, or potential safety concerns for participants.

Protocol Elements per ICH E6 [399]

1. Administrative information
2. Background and rationale
3. Trial objectives and purpose
4. Trial design
5. Selection and withdrawal of participants
6. Treatments
7. Assessment of efficacy
8. Assessment of safety
9. Statistical considerations
10. Direct access to source data
11. Quality control and assurance
12. Ethics
13. Data handling and record keeping
14. Financing and insurance
15. Publication policy

A protocol deviation is any change, divergence, or departure from the study design or procedures as defined in the study protocol. How protocol deviations are defined and processed tends to be company-specific. In general, protocol deviations are defined as major or minor. Major protocol deviations are deviations that affect the participant's safety, rights, or welfare, that affect the primary endpoint, or violate key inclusion or exclusion criteria. Minor deviations do not compromise participant welfare or safety and are minor things, such as missing a noncritical laboratory test or collecting samples outside the predefined visit window. Major and minor protocol deviations are sometimes called Important Deviations and Non-Important Deviations, respectively.

All protocol deviations must be documented, clearly explaining what happened, why it happened, the impact of the deviation, and any corrective actions taken. The 'why it happened' is sometimes called a root cause analysis: was the deviation due to human error, was it avoidable, was it a process issue, or was it due to unclear instructions, etc.? Any corrections that were taken are sometimes called Corrective and Preventive Actions (CAPA). CAPAs are put in place to prevent future deviations for the same reason. They may involve further training at a clinical site, improving the design of the Case Report Form, or clarifying the protocol text.

Protocols are not just scientific plans; they are legally binding documents during inspections by regulators as they assess protocol compliance, deviations, violations, and the impact on participant safety.

Study Objectives. Study objectives define the purpose and goals of the study. They are often broken down into primary, secondary, and exploratory. Primary endpoints are the main objectives of the study; usually, there is only one, but it is not uncommon to see more than one. The number of participants in a study is based on the primary endpoint. Secondary endpoints are supporting endpoints to the main endpoints, and there may be many. These may include quality-of-life endpoints, safety measures, pharmacokinetic assessments, or changes in biomarkers. Exploratory endpoints are hypothesis-generating, such as the effect of genotype or other subgroup analyses on an endpoint, or the examination of pharmacokinetic-pharmacodynamic relationships.

Treatments. Treatments are the interventions being compared. This can include drugs, medical procedures, behavioral therapies, a placebo, or any combination of these. For active drug treatments, the active drug is sometimes referred to as the investigational product. The group each treatment is assigned to is the treatment group. Each treatment group must be precisely defined and sufficiently discrete so that a difference in treatment effects can be detected and precisely estimated. Usually, for drug treatments, the generic name (and sometimes brand names), formulation, strength, route of administration, dose, dosing regimen, and duration of treatment must be detailed. If there are any background drugs that each treatment group receives during the treatment period, these too must be described in detail. The comparator group against which the active drug or intervention is compared can be a placebo, another drug or drugs, the standard of care, or a different formulation of the same drug.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Study Population. The study population is defined by the inclusion/exclusion criteria, which was introduced in the Introduction section of the Special Populations chapter.

Blinding. Blinding, sometimes called masking, is used to minimize bias introduced by a participant, an investigator, or a company employee into a clinical trial and refers to whether a participant knows which treatment they were assigned to. There are many different types of blinds:

- In an open-label study, everyone involved in the study knows the treatment assignment of a participant. This is often used in early-phase trials and when it is impossible to blind the study. For example, in a clinical trial comparing a new drug to a placebo, the experimental drug may cause nausea after administration. This study would be difficult to blind because participants who did not experience nausea would know they were in the placebo group. Such studies, while cheaper and easier to run, may be subject to bias, either intentional or unintentional.
- In a single-blind study, only the participant does not know which treatment they are assigned. Such a design may help reduce any placebo effect or expectation bias among participants. Still, there may be residual bias on the part of the investigator depending on how subjective the outcome assessments are. This type of blind requires all treatments to have identical-looking formulations, which may be hard to do.
- In a double-blind study, neither the participants nor the investigators know which treatment a participant is assigned to. Double-blind studies are the gold standard for pivotal trials because they protect against both participant and investigator bias. This type of blind is more complex because, like a single-blind study, it requires all treatments to have identical formulations.
- In a triple-blind study, the participants, the investigators, and the analysts/statisticians who analyze the data are blinded as to participant treatment.
- In a quadruple-blind study, the participants, the investigators, the analysts/statisticians who analyze the data, and anyone who assesses study outcomes or endpoints, like radiologists assessing tumor size, are blinded as to participant treatment. Such a design is rare.

Sometimes it is impossible to blind a treatment, such as when one treatment is drug therapy and the other is surgery, or when there are two different drugs and one has a distinguishing feature, such as a bad taste or a side effect, that gives away which treatment it is. One method to create a blind in this instance is a double-dummy blind, in which participants receive both "treatments" simultaneously. Suppose one treatment is a tablet and the other treatment is an injectable version of a drug. A double-blind study would involve administering active treatment and placebo injection simultaneously in one group, and placebo tablet and active drug injection simultaneously in the other group.

During a study, it may be necessary to unblind a participant's treatment, or even the entire study. If a participant has a serious adverse event (SEA), a serious unexpected serious adverse reaction (SUSAR), or requires medical treatment, it may be necessary to determine which treatment they received. If the DSMB determines a high rate of adverse events, mortality, or toxicity, it may be necessary to unblind the entire study. Unblinding of any kind during a study must be documented, giving the reasons, when, how, and who was unblinded. After a study is completed, all those involved in the study, participants, investigators, and analysts, will be unblinded to allow for proper statistical analysis by treatment groups.

Study Assessments and Schedule of Events. Within each study, many activities occur: serial blood samples are collected for safety laboratory tests, such as measurement of clinical chemistry and hematological values, blood samples will be collected for measurement of blood, plasma, or serum drug concentrations, urine will be collected for urinalysis and measurement of urinary drug concentrations, ECGs may be collected, blood pressure may be assessed, and assessment of the subject's state of being ("How are you feeling"?) is made. These assessments are designed to ensure participant safety and to address the specific questions the study aims to answer. Also, the primary and secondary endpoints will be measured. These activities may be done multiple times throughout the study during scheduled study visits. In the event of an adverse event, unscheduled assessments,

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

such as safety laboratories, may be conducted to ensure the participant's safety.

A key issue is whether the study assessments will be conducted by a single or centralized laboratory, which is useful when reproducibility and consistency are concerns. Using a single laboratory is more complex and costly because it often requires shipping of samples to the laboratory.

Randomization. Randomization is the process of assigning participants to a treatment or a treatment sequence, as in a crossover design. Each distinct pathway (treatment) a participant may receive is called a treatment arm. Randomization ensures approximately equal distribution of prognostic variables across treatment groups. For instance, suppose age affects disease progression. It would not be good if one treatment group had more elderly participants than the other. Randomization attempts to place approximately equal numbers of the elderly into both groups. Randomization also helps prevent investigator bias in assigning participants to treatments.

There are many ways to randomize participants to treatments; the easiest is simple randomization, which is like flipping a coin to determine which treatment a participant is assigned to. Usually, participants are assigned 1:1 in a two-arm trial, with an equal number assigned to each treatment. Still, there may be situations where other ratios are used, such as 2:1 active treatment to placebo, to minimize the number of placebo-treated participants. Block randomization was discussed earlier in the chapter and is used to ensure better balance across prognostic variables. There is also adaptive randomization, in which allocation ratios are adjusted based on interim analyses to favor a particular study arm, such as assigning more participants to the superior treatment to minimize treatment failures. In doing so, however, this results in unequal numbers of participants across treatment arms, which may affect the statistical analysis later.

The study statistician often generates the randomization scheme, and care is taken to keep the schedule secret until the end of the study.

Sample Size, Statistical Power, and Assurance. An important question for any clinical trial is how many participants should be enrolled and how they should be allocated to treatments. In a typical statistical analysis of an endpoint from a clinical trial, the null hypothesis is the hypothesis of no effect:

$$\text{Ho: mean drug effect} = \text{mean control effect}$$

and the alternative hypothesis is that the means are not equal:

$$\text{Ha: mean drug effect} \neq \text{mean control effect}$$

Our goal is to reject the null hypothesis. You can never prove the null hypothesis – it can only be disproven because statistical testing is designed to detect evidence against the null hypothesis, not to confirm it. You might have heard the phrase "absence of evidence does not mean evidence of absence." This is hypothesis testing in a nutshell. Just because we don't see it, doesn't mean it's not there. Think of it this way. Suppose we look at a pond near our house, and we see white swans. Maybe we think, "all swans are white." That becomes our null hypothesis. Our alternative hypothesis is that there is at least one non-white swan somewhere in the world. More white swans fly into the pond, and we gain further confidence in our null hypothesis. If 10,000 white swans fly into the pond, we still cannot prove the null hypothesis, because all it will take is one black swan to reject it. No amount of white swans can provide it beyond any doubt.

To test the null hypothesis, we compute our test statistic (our estimate) and, under the null hypothesis, calculate the probability of observing our estimate or a value larger than it. If the probability is less than some critical value, called alpha (α), we reject the null hypothesis and declare a difference between our drug and the control. Alpha is called the Type I error and is the probability of incorrectly rejecting the null hypothesis and declaring a difference when, in fact, there is none. For example, suppose we are measuring systolic blood pressure (SBP) in 15 participants in the drug group and 15 patients in the placebo control group. After 12 weeks of treatment with our new drug, we observed a mean blood pressure of 123 mm Hg, whereas the control group had a mean of 128 mm Hg. Assuming the variance of our measurements is the same for both drug and control,

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

and is equal to 5 mm Hg, we can use what is called a T-test to calculate the probability that the mean SBP in the drug group (123 mm Hg) equals the mean SBP in the control group (128 mm Hg). After doing the calculation, we get a p-value of 0.0106. If we had set alpha to 0.05, we would have rejected the null hypothesis; if we had set alpha to 0.01, we would not have rejected the null. This example illustrates something called p-hacking. Since most journals don't publish nonsignificant results, it's tempting not to use 0.05 as our alpha, so that we can declare our drug has an effect over placebo.

This is where the concept of statistical power comes into play. First, let's define a treatment effect, which is the change in an outcome between the treatment of interest and the comparator group. For example, this can be the difference in means, the difference in risk, or the difference in the odds of some event occurring, and it can be expressed as a difference or a ratio, e.g., an odds ratio. Power is the probability of correctly detecting a treatment effect, or the probability of rejecting the null hypothesis when the alternative hypothesis is true. Several factors influence the power of a test:

- Sample size: The larger the sample size, the larger the power to detect a treatment effect.
- Effect size: How big is the difference between the treatment and control? The larger the effect size, the larger the power. The effect size is generally estimated as the difference in the means between the treatment of interest and the comparator divided by the common standard deviation of the observations.
- Variability: the smaller the variability, the larger the power to detect a treatment effect.

Of these, sample size and variability can be manipulated by researchers. A more precise measuring instrument could improve the power of a statistical test by reducing variability. More often, however, the sample size is chosen so that the power to detect a difference is set to a predetermined probability, such as 80% or 90%. For example, suppose we wish to detect a 5 mm Hg difference in systolic blood pressure (SBP) for a new drug with a 90% chance of detecting a treatment effect if it truly exists. To calculate the sample size, the researcher must estimate the effect size. If the standard deviation of SBP in our hypertensive patient population is 18 mm Hg, and using a 2-sided Student's T-test with alpha set to 0.05, we would need ~190 patients per treatment arm. Assuming 10% of participants may drop out of the study, the total sample size would then be about ~420 participants. In summary, we may write "Detecting a 5 mmHg difference in SBP with 90% power and a standard deviation of 15 mmHg, assuming 10% dropout, requires ~420 subjects total."

Suppose we weren't so sure about the 5 mmHg difference between treatments. Suppose we run the trial and find that the treatment difference is 3 mmHg; this would decrease our probability of detecting a treatment difference and would be a bad situation. Assurance is a weighted average of power that accounts for the likelihood of other scenarios by averaging power across all possible outcomes. In this manner, this can design a more robust trial. What assurance does is essentially account for the uncertainty in the effect size.

Monitoring. Clinical trials are monitored throughout by professionals who examine the raw data collected at study sites to ensure data quality and prevent fraud. These professionals go by many names, including study monitors and clinical trial associates. They review all documentation, such as signed Informed Consent forms and training logs, to ensure patient safety and that clinic staff are following the protocol. It is often impossible to monitor every piece of information in large-scale clinical trials, and today, most companies use a risk-based approach that focuses on critical data and risks important to patient safety, rather than using 100% data monitoring.

A variety of monitoring levels is often used. At the base level, monitoring is done on-site by individuals. Monitoring is also done off-site, aggregating across individuals to identify outliers, errors, and fraud. The future of monitoring will involve a "boots on the ground" approach and likely using machine learning to identify errors in the data.

You might think that fraud is unlikely to occur in a clinical trial, but I suppose it depends on how you define 'unlikely.' In a survey of 549 authors who published clinical trials between 1998 and 2001, 5% reported fabrication or misrepresentation in a study they participated in within the last 10 years, and

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

17% personally knew of cases or misrepresentation in the last 10 years [400]. Overall, the incidence of data fraud has been around <1% [401]. One particularly noteworthy case of fraud involved Ketek (telithromycin), an antibiotic approved in 2004 for the treatment of community-acquired respiratory infections [402]. In 2001, before approval, Sanofi initiated Study 3014, a study designed to compare the rates of hepatic, cardiac, and visual adverse events in telithromycin-treated patients with those in patients treated with amoxicillin-clavulanate. More than 1,800 doctors participated in the clinical trial, many of whom were new to clinical research, and were paid up to $400 per patient enrolled. More than 24,000 patients were recruited. In a review of the data during monitoring, the doctor who enrolled the most patients, more than 400, which included her entire office staff, Dr. Maria "Anne" Kirkman-Campbell, a family physician from Gadsden, Alabama, was found to have committed fraud, including the fabrication of patients; 91% of the patients she enrolled were fabricated. Other sites were investigated, and in all, four of 10 sites were found to have committed fraud and were referred for criminal investigation. Dr. Kirkman-Campbell was sentenced to 57 months in prison and ordered to pay $557,000 in fines and $925,000 in restitution. Despite these data quality issues, Ketek was approved in 2004 but was removed from the market in 2007 due to adverse events. Sanofi-Aventis was ordered to appear before Congress to testify on what they knew and when. In a major scandal, one of the committee members claimed that "only a company intent upon ignoring the obvious could have failed to detect fraud in Study 3014." [403] No company wants to be embarrassed like that again.

Operational Considerations. Once a protocol is written, it must be executed. Some studies can be conducted at a single site, but others may be so complex that they require multiple sites across multiple countries. This can lead to challenges with the consistency and reliability of results across sites and among different investigators at each site. Site selection is a major concern for companies running studies, as it takes time and money to enroll a site. Low-performing sites may be a waste of time and money. Further, high-performing sites that recruit many participants but lack consistency and do not follow the protocol can also lead to data quality issues. For this reason, all sites and investigators are trained on the protocol and its procedures before enrolling any patients. Also, clinical sites are monitored throughout the study to ensure data quality.

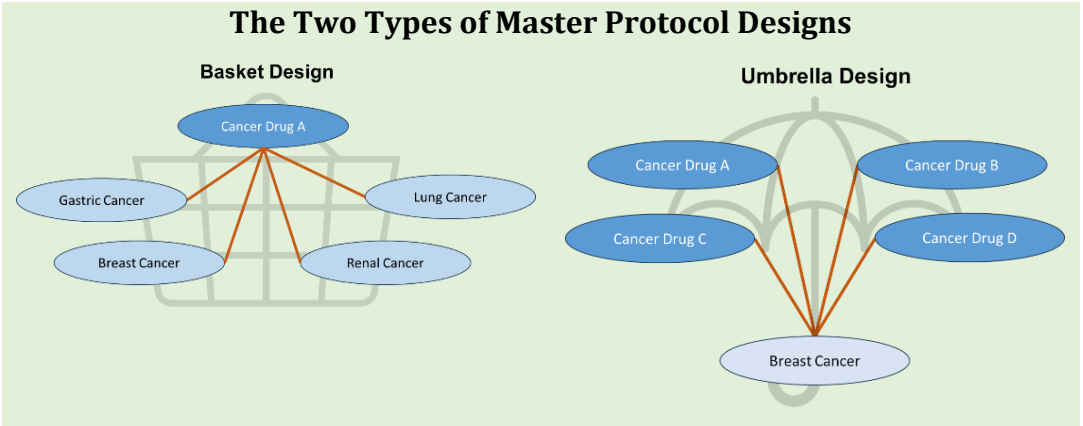
There is also the issue of getting the drug and study supplies to all the sites. Shipping experimental drugs into different countries may require different regulatory approvals. Once at the site, the pharmacy may need to be trained on how to store and prepare the drug. Inadequate storage may cause the drug to degrade and not be safe for administration to humans. All drugs sent to a site must be accounted for. When the study is over, any excess drug remaining must be properly disposed of.

There is also the issue of getting data from the site to the sponsor. Clinical sites must complete Case Report Forms (CRFs), whether paper or electronic, that capture all protocol-related data for each participant. CRFs are the primary interface between all site data and the trial database. If there is any issue with some piece of data, such as an unlikely event, the CRF will be reviewed to ensure it was transferred correctly to the electronic database, or the source data may be reviewed to ensure the results were transferred correctly to the CRF. Clinical Trial Monitors ensure that data recorded on CRFs are accurate. All CRFs must follow ALCOA principles: Attributable, Legible, Contemporaneous, Original, Accurate, plus Complete, Consistent, Enduring, and Available. Typical sections include: demographics, informed consent confirmations, inclusion/exclusion criteria checks, medical history, concomitant medications, date and time of study treatment administration, efficacy assessments, safety data, protocol deviations, and study completion or discontinuation details.

Recent Advances in Clinical Trial Design and Development

The drug development paradigm presented in this chapter has been largely considered a traditional development program. Clinical development of drugs through this pathway can be long, expensive, and inconvenient for participants who have to go to a site for their visits and monitoring. Companies are now looking at ways to reduce clinical development costs and timelines, and make their studies more participant-friendly. One way they are doing this is by improving the design and conduct of clinical trials. This is such an important topic that, in 2024, the FDA created a new center, the CDER

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Center for Clinical Trial Innovation (C3TI), focused on Clinical Trial Innovation. A few of the more impactful advances in clinical trial designs will be discussed in this section.

Seamless Designs. Sometimes sponsors combine Phase 2 and Phase 3 into a single seamless study. [404]. In this design, in Phase 2, the learning stage, multiple doses or treatment combinations will be studied. Treatment combinations may be added or discontinued, and dose optimization may be performed. In the confirmatory stage, Phase 3, the efficacy and safety of a particular treatment are studied. For seamless designs, results from Phases 2 and 3 are used to evaluate both the final safety and efficacy of the treatment. Seamless designs are quite common in oncology but less common in other areas.

Master Protocols. In oncology, it is not immediately clear which cancer types a drug may work in, so sponsors will conduct small, individual pilot studies across many indications to determine which cancer type has the best response rate. This is inefficient. A Master Protocol is a single study that includes multiple groups or mini-studies with the same or different diseases, using the same or different drugs to treat them under a single protocol, making it a more efficient way to conduct drug development [405]. There are generally two types of master protocols: basket designs and umbrella designs. In a basket design, one drug is studied across multiple indications, whereas in an umbrella design, multiple drugs are tested in a single indication. A lot has been said about Gleevec (imatinib), which was the first targeted cancer therapeutic approved for use. But imatinib also has another first. In 2006, when imatinib was approved, it was approved for five indications based on a single Phase 2 clinical trial. This stunned the pharmaceutical community – how could Novartis get so many indications at once, since the FDA usually only approved a drug for a single cancer type? They did so using a basket design, testing multiple tumor types simultaneously. This design saved Novartis time and costs and provided benefits to participants, as indications were approved faster than usual. For more information on Master Protocols, the FDA issued Guidance to Industry in 2020 [406].

Another type of master protocol worth discussing is the platform trial, which is similar to umbrella trials in that it studies a single disease. However, the key difference is that participants are randomized to a treatment at the beginning of the trial based on the presence or absence of a specific biomarker. As the trial progresses, treatment arms may be stopped or added as information accumulates. One of the longest-running adaptive platform trials is the I-SPY 2 trial (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2, NCT01042379) for patients with Stage 2-3 breast cancer with high risk for recurrence. Before enrollment, each patient's breast cancer is categorized into one of 10 different subtypes using hormone receptor status, HER2 status, and MammaPrint® scores, which test 70 different genes associated with early breast cancer and characterize them as either high-risk or low-risk. Participants are then adaptively randomized based on their subtype to a control arm or one of the experimental arms. Since its inception in 2010, over 3000 participants have been treated, and 12 treatment combinations have completed accrual, of which seven have progressed into one breast cancer tumor

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

subtype, indicating an 85% chance that the combination will be successful in confirmatory Phase 3 studies. Two of these seven drugs have already received accelerated approval (Keytruda and Perjeta), and one has received breakthrough designation (trilaciclib). One of the many unique things about I-SPY is that this is an academically sponsored protocol that uses experimental drugs from different companies.

Decentralized Trials. In a traditional clinical trial, participants enroll in a study. At various times, they must go to a central site, usually a hospital, where they might be dosed with a drug, have samples collected, and have their vitals checked and be queried about medical events that occurred since the last visit. Traveling to a central site can be onerous, and in rural areas it may involve driving hundreds of miles. Decentralized studies aim to shift activities away from a central site toward the participant, allowing study activities to be conducted in non-traditional settings, such as the participant's home or a local pharmacy, or by moving from a hospital to their local doctor. Event monitoring may be performed remotely using wearables, ingestible devices, or implantable devices. Blood samples may be collected by the participant using new blood collection devices. Microsoft Teams or Zoom may be used to talk with the participant and obtain adverse event information or to obtain informed consent. It should be stressed that the vigor and quality expected for a centralized site still apply to decentralized sites. Decentralized trials are more patient-centric, a buzzword in the industry for healthcare that is more patient-centered, and patients actively participate in their medical treatment. Decentralized trials may improve recruitment by making participation easier and, by reducing barriers to entry, make studies more inclusive. An FDA Guidance on this topic discusses in greater detail how to design and implement decentralized trials [407].

Statistical Analysis and Reporting of Pivotal Trials

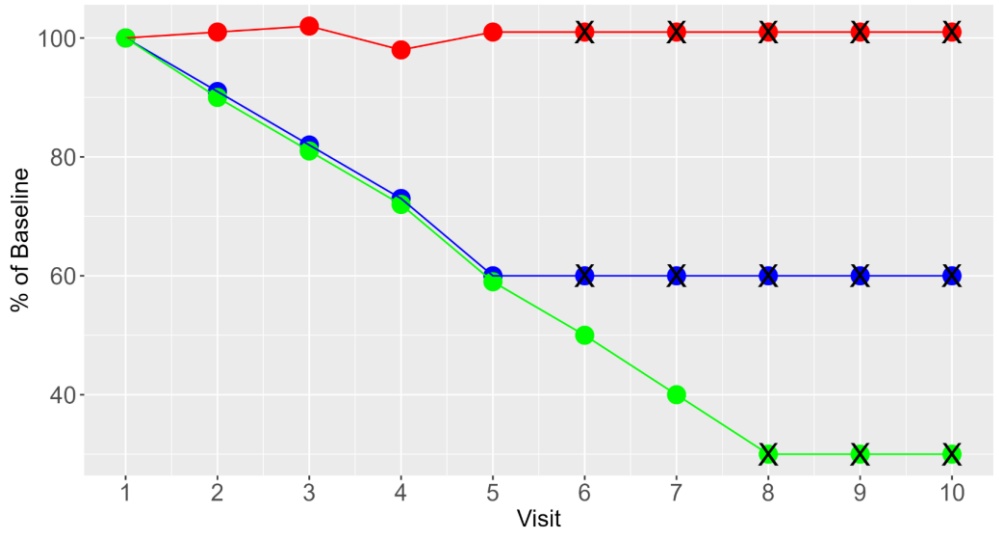
Participant Disposition. Part of any analysis of a clinical trial is patient disposition, as it shows who actually contributed to the study results and subsequent data analysis. How many were screened? How many screen failures were there? How many were randomized to treatment?> How many received treatment? How many completed the study? How many withdrew? How many were lost to follow-up? How many were included in the primary analysis? Regulators require a complete disposition of each participant.

Understanding participants' disposition is important for assessing bias and the validity of the trial. For example, a high dropout rate in a particular treatment arm may invalidate the results of a trial. Suppose in a trial of two different drug doses and a placebo, the dropout rate was higher than that of the placebo in the lowest active drug group and highest in the highest dose group. This could have been due to a dose-related increase in adverse events. The treatment effects in each group then reflect only those patients who did not drop out and may not be generalizable to the population at large.

Demographics. Baseline demographic and clinical characteristics for each treatment group must be displayed to ensure similarity between groups. The demographics of pivotal clinical studies should reflect the population in which the drug will be used. It is not necessary to perform a statistical analysis to demonstrate equality among groups. Typically, age, weight, sex, race, baseline disease severity, and other important prognostic factors are reported, along with the sample sizes used to calculate each estimate. For continuous variables, the mean, standard deviation, and range are reported, whereas for categorical variables, the number and percent in each category are presented.

Missing Data. Not all participants who enroll in a clinical trial will complete it, and even if they do, not all participants will complete all the required assessments. Such data are said to be 'missing.' Missing outcome data are expected in clinical trials, and their presence does not necessarily invalidate the study's results. However, how such missing outcome data are handled can lead to bias in the study results and interpretation. What makes this issue so challenging is that there are many ways to handle missing data, both for outcomes and other variables, none of which are universally accepted or applicable [408].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026



Appendix 2 Figure 3: Illustration of the effect of LOCF on treatment estimation in the presence of missing data due to dropout. Suppose there are three participants to be assessed for the clinical endpoint across 10 visits. Subject 1 (denoted in red) and Subject 2 (denoted in blue) drop out of the study after Visit 5, whereas Subject 3 (denoted in green) drops out after Visit 7. LOCF values are denoted as symbols overlaid with an X. In Subject 1 (denoted in red), the drug has no effect, and LOCF does a reasonable job of estimating the treatment effect. The drug has almost the exact same effect in Subject 2 and Subject 3. But because Subject 3 drops out after Visit 5, the estimate of the treatment effect is overestimated, since had they stayed in the study, their results would have more closely mirrored Subject 3's. Hence, because LOCF freezes the last observation, it overestimates the outcome for participants who drop out early.

To understand how to handle missing data, it is first necessary to identify the reason for missingness. If data are missing and the reason has nothing to do with the data itself, either observed or unobserved, then the data are said to be Missing Completely at Random (MCAR). For instance, a biomarker cannot be assessed because the sample broke in the centrifuge before it could be analyzed, or the participant could not make it to the site for a visit because their car broke down; these are completely random events. Missingness is completely accidental with MCAR. If, however, missingness is systematic and the reason for missingness depends on the participant's characteristics, and not related to the outcome itself, this is said to be Missing at Random (MAR). MAR is a weaker assumption of MCAR, but it is often more realistic. For example, older participants may be more likely to miss a visit, or participants with higher baselines may be more likely to miss a visit. In this case, age and baseline are the reasons for missingness, and these data are MAR. The last reason for missingness is if the missingness depends on the thing being measured. For example, suppose in a depression trial, patients who are very depressed are more likely to miss a visit, or that patients who do not respond in a weight loss trial drop out of the study because the drug is not working. In this case, the missingness is nonignorable; you cannot ignore it. It is important to understand the reason for missingness because the statistical methods that handle missing data often depend on it.

One of the easiest ways to deal with missing data is to ignore it and analyze only subjects with complete data. This is known as complete case analysis (CCA). CCA is a valid method when the data are MCAR. CCA can yield biased results when the data are MAR or nonignorable. Many studies using antidepressants or pain medications in the 1990s and 2000s were analyzed using CCA despite there being a 30 to 40% dropout, which would be considered nonignorable because dropout was due to adverse events or lack of efficacy. Many of these studies were found to be statistically significant. Reanalysis using more appropriate statistical models, appropriately accounting for the missing data,

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

later showed that these studies were, in fact, not statistically significant [409].

Other options to analyze outcome data that include missing values are to use statistical tests that don't depend on all data being present and are MAR, like mixed effect models, or to try to estimate the values of the missing data, i.e., impute the missing value, and then analyze the data, including all observed data and all the imputed data, as if it were not missing. A simple, and often incorrect, way to handle missing data due to study dropout is to carry the last observation forward in time to all remaining visits. This is called the Last Observation Carried Forward (LOCF) method (Appendix 2 Figure 3). An example of LOCF and how it can lead to a wrong interpretation is the CATIE trial, a multi-arm randomized controlled clinical trial of different antipsychotic drugs designed to see how they work in the real world [410]. Of the almost 1,500 participants enrolled in the trial, 64% to 82% discontinued therapy before 18 months, often due to lack of efficacy or intolerable side effects. Analysis of the symptom scale used to assess the effectiveness of antipsychotic drugs, the Positive and Negative Symptom Scale (PANSS), was performed using analysis of covariance with LOCF for missing data due to dropout. The results showed that improvement was greatest in the olanzapine cohort and least for the quetiapine cohort. Olanzapine-treated patients tended to drop out later, and their efficacy trajectory was better captured. Since LOCF freezes the last observation, it overestimates the outcome for patients who drop out early. Hence, the efficacy of the other drugs in the study is underestimated [411].

Handling missing data is complicated and the wrong method can lead to biased results. When reviewing clinical trial analyses, the handling of missing data should be examined in the methods section.

Statistical Analysis. After the trial is completed, the study's results are statistically analyzed. What constitutes a successful or failed clinical trial depends on whether the trial achieves "statistical significance," which depends on how the data were analyzed. A statistically significant clinical trial means the results were due to the treatment of interest and not to chance. The analysis of clinical trial data might seem straightforward, but it is not – far from it; it is very, very complicated. Different analytical methods can yield different trial results. Great effort is put into predefining, before looking at any data, how the data will be analyzed. Statistical methods must be appropriate, unbiased, precise, and robust to withstand the misspecification of analysis assumptions. The 1998 E9 Guidance on Statistical Principles for Clinical Trials is meant to provide companies with guidance on appropriate clinical trial design and analysis of clinical trial results [408]. This guidance is 46 pages long, making it one of the longest guidances issued by regulatory authorities. The guidance makes no specific recommendations for statistical tests, so sponsors have flexibility in how they analyze the data. Companies must prepare *a priori* a Statistical Analysis Plan that details the population to be analyzed in the clinical trial, the statistical methods, how statistical significance will be assessed, how missing data will be handled, and other relevant procedures. The FDA generally reviews all Statistical Analysis Plans, and no analyses are done until both the sponsor and FDA agree on the statistical methods. It is beyond the scope of this book to discuss all the issues related to statistical analysis of clinical trial data. Still, a few topics are worth discussing because they are reported in the package insert and on regulatory sites.

Clinicaltrials.gov is a useful website for patients looking to enroll in clinical trials. It is generally user-friendly. Over the years, web managers have required sponsors to tone down the scientific language, making it easier for non-scientists to read. Lately, however, sponsors have been posting trials that include some new terminology that will not be immediately apparent. For example, Regeneron posted a Phase 3 clinical trial (NCT03399786) evaluating the efficacy and safety of evinacumab, a monoclonal antibody, for the treatment of patients with familial homozygous hypercholesterolemia. The primary measure of outcome was defined as: "*Percent change in calculated low-density lipoprotein cholesterol (LDL-C) from baseline to week 24 (Intent-to-Treat [ITT] Estimand).*" Most of that sentence makes sense, except for that last part. What is an estimand? Right now, my Microsoft Word dictionary doesn't recognize it as a valid word. A quick check of Merriam-Webster's Online dictionary (as of 2025) shows that no such word exists in its database either.

The primary regulatory guidance to support the statistical analysis of clinical trials is ICH E9 –

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Guideline on the Statistical Principles for Clinical Trials [181]. Issued in 1998, it has had no updates until 2020, when it was updated for a single topic: Estimands [412]. Estimands precisely describe the treatment effect a study intends to quantify [413]. Estimands are designed to improve clarity around what is being measured, and their description should reflect the clinical question of interest, accounting for factors that might affect the outcome, known as intercurrent events. The statistical analysis of the clinical trial should then reflect the estimand. Intercurrent events, another poorly chosen term by statisticians, include patients who drop out of a trial before completing it, patients who do not take their medication as directed, or patients who may start other treatments while on study. The estimator is the statistical method used to estimate the treatment effect as defined by the estimand, and the estimate is the actual value computed by the statistical test.

Kahan et al. provide an example of this framework [413]. Dupixent (dupilimab), a monoclonal antibody used to treat chronic obstructive pulmonary disease, was compared with placebo after 12 weeks of treatment in patients with moderate-to-severe asthma, using the maximal amount of air a person could exhale from their lungs in 1 min, called FEV1 [414]. Some patients stopped taking the drug because it didn't work, and/or they also took rescue medications like bronchodilators to treat disease exacerbations. The estimand was then *"the difference in the mean FEV1 at Week 12 between dupilimab and standard of care vs. placebo and standard of care, in patients with uncontrolled asthma, if they were to continue using dupilimab throughout the entire trial without the use of rescue medication."* The estimator was defined using a linear mixed-effects model (which is not crucial to understanding the results). All data after the event were treated as missing in the statistical analysis. The statistical analysis showed that FEV1 increased by 0.14 liters ($p < 0.001$) compared with placebo at Week 12; this is the estimate. In this case, dupilimab improved lung function compared with placebo, accounting for intercurrent events that could have jeopardized the study's validity.

If my p-value was 0.03, this does NOT mean:

- X There is a 97% probability that the treatment works
- X The null hypothesis has a 3% chance of being true
- X The result will replicate with 97% probability

The result of a statistical analysis is often reported as a p-value, e.g., $p < 0.001$. When analyzing data to determine if a statistical analysis is "significant", two hypotheses are made. One is that there is no difference between the treatment and control groups; this is referred to as the null hypothesis. The other is that there is a difference between the treatment and control group; this is called the alternative hypothesis. Only one of these hypotheses can be correct. Statistical analysis is designed to disprove the null hypothesis. The null hypothesis cannot be proven; it is only rejected. Rejection of the null hypothesis is made if the probability of the null hypothesis being true is small. What is a 'small' probability? This has generally been defined as the α-level and is often taken to be 0.05 or 0.01, i.e., there is only a 5% or 1% chance of making a 'Type I' error – wrongly concluding that a difference exists between groups when, in fact, there is none. A p-value less than the α-level value indicates that the results were statistically significant.

P-values are frequently misinterpreted as the probability that the null hypothesis is true, or that small p-values are evidence that the effect of the treatment is large, or that it is the probability of committing a Type I error. All of these are wrong! A p-value is the probability of observing a test statistic or estimate as large or larger than what was observed ***if the null hypothesis were true***[15]. Suppose a drug trial were run and the p-value was 0.02. This would be interpreted as: if the drug had no effect, you would see results this extreme (or more extreme) only 2% of the time, and you would reject the null hypothesis of no effect at the 0.05 level. Conducting two clinical trials decreases the probability of an overall Type I error. Suppose a second clinical trial had a p-value of 0.03. In that case, the overall probability of making a mistake and concluding the drug is effective is $0.02 \times 0.03 = 0.006$, so there is a 6 in 1000 chance to erroneously conclude the drug was effective when, in fact, it was not. This is why the original 1968 FFDCA required two pivotal trials to reduce

[15] I am underlining, bolding, and italicizing this to really show how important this part of the definition is.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

the likelihood of approving an ineffective drug.

P-values are frequently criticized because they depend on sample size, i.e., the number of subjects being analyzed. Larger sample sizes decrease the error in estimating an estimand, such as the mean difference between two groups, and decrease their associated p-values. Two studies measuring the same endpoint can have the same p-value but vastly different estimates of the test statistic, due to differences in sample size. For this reason, there has been a shift in the reporting of clinical trials to not only provide p-values but also confidence intervals[16] around the test statistic. For example, the mean difference in weight loss between placebo and a new drug after 6 months of treatment may be 30 kg with a 95% confidence interval of 20 to 40 kg. Confidence intervals assess the precision of the estimate of the estimand. Larger confidence intervals mean less precision and confidence in the estimate. Suppose the confidence interval contains zero, e.g., -5 to 10 kg. In that case, the null hypothesis of no difference cannot be rejected because we cannot be certain our estimate is different than zero.

Statistical analyses of clinical trial data must be predefined and analyzed as planned after the trial is completed. Analyzing the data after the trial is complete and having seen the data is not an acceptable practice because statistical tests can be manipulated to obtain results that are 'statistically significant.' This has been called p-hacking, p-mining, or p-fishing. One of the most significant criticisms in science currently is the reproducibility crisis, in which published results cannot be replicated or are inconsistent when new data are obtained. P-hacking has been blamed as one of the reasons for this crisis.

Lastly, two important terms in the statistical analysis of clinical trials are intent-to-treat (ITT) and per-protocol (PP), which refer to the criteria for including patients in the statistical analysis. ITT analyzes all patients who were randomized to treatment, regardless of whether they ever took the drug, if they took the wrong drug, or if they did not take the drug as directed. It is intended to represent the real world and provides a more accurate estimate of a treatment's effectiveness. PP analyzes only those patients who followed the protocol's directives and took the drug as prescribed. PP may lead to imbalance across treatment groups, resulting in a biased estimate of the treatment effect. ITT provides a more conservative estimate of the treatment effect than PP.

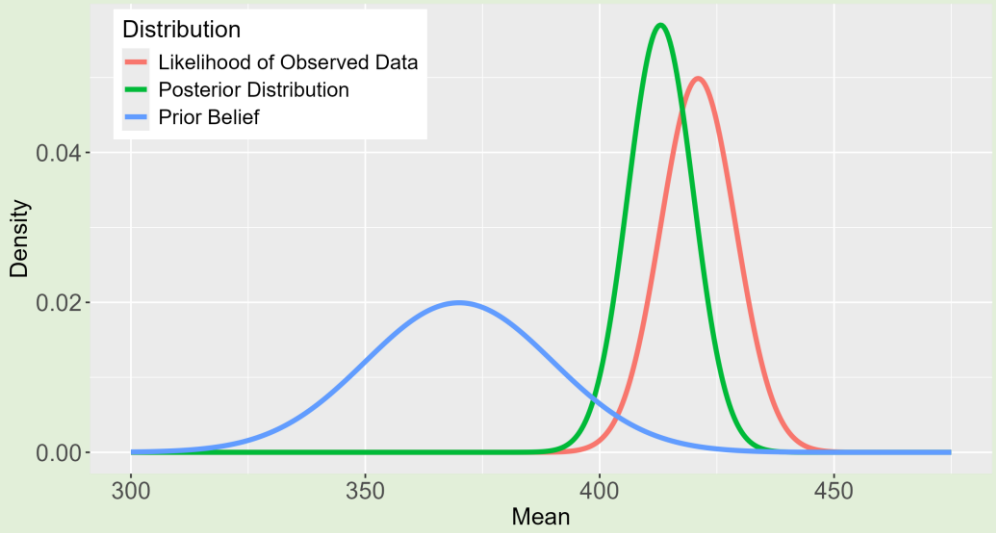
When the study is complete, the results should be reported to regulatory authorities and clinicaltrials.gov. Often, the results are published in major clinical journals, such as JAMA, the New England Journal of Medicine, or the BMJ. Like protocols, the reporting of clinical trials varies by study and company. There are some core elements, such as the reporting of estimates and p-values, but many reports and publications have been criticized for providing inadequate details. Hence, like Spirit guidelines for reporting of protocols, reporting guidelines like the CONSORT (Consolidated Standards of Reporting Trials) statement have been issued [415].

Bayesian Statistics. There are two different statistical philosophies: frequentist and Bayesian. Frequentists have a long-term interpretation of results. Frequentists interpret p-values as the long-run relative frequency of an event under repeated identical experiments. For example, if I had a p-value of 0.03 and could repeat the experiment 100 times, then three of the trials would have an estimator as large or larger than the one I observed in this trial. Bayesians interpret p-values as the probability of the null hypothesis. For a frequentist, what has happened in the past does not influence the future, and any statistics or interpretation of the data is based only on what was collected in the experiment. For a Bayesian, the past influences the interpretation of new data, and any statistics or interpretations of the data are a kind of weighted average (the posterior) of what has happened in the past (the prior) and in the experiment (the likelihood). There is a great debate between the camps over which philosophy is most relevant, and I have taken some liberties herein for ease of exposition.

[16] I don't even want to try to explain a confidence interval because if I did you would say "What the hell, that doesn't make sense!" True, it doesn't, at least not in a frequentist sense. Confidence intervals are frequently misinterpreted as the probability that the true value lies within the interval, but this is wrong; only Bayesian confidence intervals can be interpreted in this manner.

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Bayesian Statistics



Bayesian statistics is a statistical philosophy that treats unknown parameters, such as the mean, as random variables, and the probability distribution of that parameter depends on both prior expectations and observed data. In contrast, frequentist statistics treat parameters as fixed and base inference solely on the observed data and its sampling distribution.

For example, suppose we are interested in estimating the mean of a variable, and we think it has a mean of 370 and a standard error of 20 (blue line). We call this our prior distribution. Now we go out and collect data, and we observe a mean value of 421 with a standard error of 8 (red line). This is called our likelihood, and 421 would be our frequentist maximum likelihood estimate of the mean. Under a normal-normal Bayesian model, our Bayesian estimate is a kind of weighted average of the prior and likelihood and, in this case, would have a mean of 413 with a standard error of 7 (green line). This is called our posterior distribution. Bayesian statistics is often criticized for yielding different posterior distributions under different priors and is, hence, subjective by nature. Critics argue that the prior injects personal belief into scientific inference.

Besides the computational complexity of Bayesian analysis, another controversial aspect of Bayesian statistics is the choice of prior distribution, i.e., what you believe before collecting the data. The result of any Bayesian analysis is a distribution of outcomes, called the posterior distribution, that depends on the prior distribution and the likelihood of the observed data. If we have strong feelings about the prior distribution, that might unduly influence the posterior distribution of our estimate, and conversely, if we don't have strong feelings, and use what is called a non-informative prior, the prior may not affect the posterior distribution. Further, the prior is not fixed; different people may have different priors, and hence, different people's priors may lead to different posterior distributions. For this reason, Bayesian statistics is criticized as being too subjective.

ICH E9 does not choose between the two philosophies for analyzing clinical trial data; instead, it focuses on frequentist methods because they are predominant in the clinical trial setting. To date, few pharmaceutical clinical trials have used Bayesian statistical analysis of the primary endpoint as the sole analysis method; more medical devices have been approved using Bayesian statistical analysis of primary endpoints [416]. FDA guidance on the use of Bayesian statistics in medical device trials has been issued [417], and a draft guidance on their use in pharmaceutical trials has also been issued [418].

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

References

1. Feltman, R., et al. *Why pregnant people are left out of drug safety studies*. 10/24/2025 [cited 10/24/2025; Available from: https://www.scientificamerican.com/podcast/episode/the-dangerous-data-gap-in-pregnancy-drug-research/.
2. United States Code of Federal Regulations, *Title 21: Food and Drugs*. 2023 (https://www.ecfr.gov/current/title-21).
3. Sanchez, M.C., *Food Law and Regulation for Non-Lawyers: A U.S. Perspective*. 2015, Springer: Heidelberg
4. Massie, L.X. and F. Leborgne Nowels, *New FDA Regulatory Framework for Cosmetics: The Modernization of Cosmetics Regulation Act of 2022 (MoCRA)*. 2023 (https://www.akerman.com/en/perspectives/new-fda-regulatory-framework-for-cosmetics-the-modernization-of-cosmetics-regulation-act-of-2022-mocra.html).
5. Hemila, H. and E. Chalker, *Vitamin C for preventing and treating the common cold.* Cochrane Reviews, 2013. **Issue 1**: p. Art. No.: CD000980 (https://doi.org/10.1002/14651858.CD000980.pub4)
6. Wade, N., *Life-extending chemical Is found in certain red wines*, in *The New York Times*. 2003 (https://www.nytimes.com/2003/08/24/health/lifeextending-chemical-is-found-in-certain-red-wines.html).
7. Popat, R., et al., *A phase 2 study of SRT501 (resveratrol) with bortezomib for patients with relapsed and or refractory multiple myeloma.* British Journal of Haematology, 2012. **160**: p. 714–717 (https://doi.org/10.1111/bjh.12154)
8. Ichim, M.C., *The DNA-Based authentication of commercial herbal products reveals their globally widespread adulteration.* Fronteirs in Pharmacology, 2019. **10**: p. 1227 (doi: 10.3389/fphar.2019.01227)
9. Booker, A., et al., *St John's wort (Hypericum perforatum) products - an assessment of their authenticity and quality.* Phytomedicine, 2018. **40**: p. 158–164 (https://doi.org/10.1016/j.phymed.2017.12.012)
10. Booker, A., et al., *The authenticity and quality of Rhodiola rosea products.* Phytomedicine, 2016. **23**: p. 754–762 (doi: 10.1016/j.phymed.2015.10.006)
11. NOW. *NOW Reports New Testing of CoQ10 & Quercetin in Brands Sold on Amazon Shows No Improvement in Policing Quality*. 2022 09/19/2025]; Available from: https://www.nowfoods.com/healthy-living/articles/now-tests-coq10-quercetin-sold-amazon.
12. Brasky, T.M., E. White, and C.-L. Chen, *Long-term, supplemental, one-carbon metabolism–related vitamin B use in relation to lung cancer risk in the Vitamins and Lifestyle (VITAL) Cohort.* Journal of Clinical Oncology, 2017. **35**: p. 3440–3448 (https://doi.org/10.1200/JCO.2017.72.7735)
13. Geller, A.I., et al., *Emergency department visits for adverse events related to dietary supplements.* New England Journal of Medicine, 2015. **373**: p. 1531–1540 (DOI: 10.1056/NEJMsa1504267)
14. Cohen, P.A., et al., *Presence and quantity of botanical ingredients with purported performance-enhancing properties in sports supplements.* JAMA Network Open, 2023. **6**: p. e2323879 (doi:10.1001/jamanetworkopen.2023.23879)
15. Rosenberg, S.A., R. Spiess, and R. Lafreniere, *A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes.* Science, 1986. **233**: p. 1318–1321 (DOI: 10.1126/science.3489291)
16. Mitra, A., et al., *From bench to bedside: the history and progress of CAR T cell therapy.* Fronteirs in Immunology, 2023. **14**: p. 14 May 2023 (https://doi.org/10.3389/fimmu.2023.1188049)
17. McCarthy, B. *The mRNA COVID-19 vaccines are "really gene therapy."*. 31 Aug 2021; Available from: https://www.politifact.com/factchecks/2021/aug/31/joe-rogan/joe-rogan-falsely-says-mrna-vaccines-are-gene-ther/.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

18. Banoun, H., *mRNA: vaccine or gene therapy? The safety regulatory issues.* International Journal of Molecular Science, 2023. **24**: p. 10514 (doi: 10.3390/ijms241310514).
19. Washington University in St. Louis, Health Communication Research Laboratory, and iHeard STL. *Claims that mRNA vaccines change human DNA are back*. 2024; Available from: https://hcrl.wustl.edu/claims-that-mrna-vaccines-change-human-dna-are-back/.
20. Food and Drug Administration. *What is Gene Therapy?* 2018; Available from: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/what-gene-therapy.
21. Food and Drug Administration, *Office of Generic Drugs 2021 Annual Report*. 2021 (https://www.fda.gov/drugs/generic-drugs/office-generic-drugs-2021-annual-report).
22. Andrade, C., *Bioequivalence of generic drugs: a simple explanation for a US Food and Drug Administration requirement.* Journal of Clinical Psychiatry, 2015. **76**: p. e742–e744 (DOI: 10.4088/JCP.15f10094)
23. Siddiqui, Z., *India's Dr Reddy's hit by FDA warning, shares fall 15 percent*, in *Reuters*. 2015 (https://www.reuters.com/article/us-dr-reddys-fda-warning/indias-dr-reddys-hit-by-fda-warning-shares-fall-15-percent-idUSKCN0SV0FA20151106).
24. Pollack, A., *Valeant Under Investigation for Its Drug Pricing Practices*, in *NY Times*. 2015 (https://www.nytimes.com/2015/10/15/business/valeant-under-investigation-for-its-drug-pricing-practices.html).
25. Statista, *Leading 50 global pharmaceutical companies by prescription sales and R&D spending in 2022*. 2023 (https://www.statista.com/statistics/273029/top-10-pharmaceutical-companies-sales-and-rundd-spending-in-2010/).
26. Statista, *Aerospace and defense companies with the highest spending on research and development (R&D) in 2022*. 6–25–2025 (https://www.statista.com/statistics/1102967/global-research-and-development-top-aerospace-defense-spenders/).
27. Statista, *Pfizer's expenditure on research and development from 2006 to 2022; Roche's expenditure on research and development from 2011 to 2022; Johnson & Johnson's expenditure on research and development from 2005 to 2022*. 6–20–2025 (https://www.statista.com/statistics/267810/expenditure-on-research-and-development-at-pfizer-since-2006/ ; https://www.statista.com/statistics/266518/roches-expenditure-on-research-and-development-since-2007/ ; https://www.statista.com/statistics/266407/research-and-development-expenditure-of-johnson-und-johnson-since-2006/).
28. Schlander, M., et al., *How much does it cost to research and develop a new drug? A systematic review and assessment.* Pharmacoeconomics, 2021. **39**: p. 1243–1269 (DOI: 10.1007/s40273-021-01065-y)
29. Brown, D.G., et al., *Clinical development times for innovative drugs.* Nature Review Drug Discovery, 2021. **21**: p. 793–794 (DOI: 10.1038/d41573-021-00190-9)
30. Yamaguchi, S., M. Kaneko, and M. Narukawa, *Approval success rates of drug candidates based on target, action, modality, application, and their combinations.* Clinical and Translational Science, 2021. **14**: p. 1113–1122 (DOI: 10.1111/cts.12980)
31. Duijnhoven, R.G., et al., *Number of patients studied prior to approval of new medicines: a database analysis.* PLoS Medicine, 2013. **10**: p. e1001407 (DOI: 10.1371/journal.pmed.1001407)
32. Hutchinson, N., et al., *Patient participation in clinical trials of oncology drugs and biologics preceding approval by the US Food and Drug Administration.* JAMA Network Open, 2021. **4**: p. e21104546 (doi:10.1001/jamanetworkopen.2021.10456)
33. Ball, C. and R.N. Westhorpe, *The history of simple analgesics.* Anesthesia and Critical Care, 2011. **39**: p. article 331 (DOI: 10.1177/0310057X1103900301)
34. Jeffreys, D., *Aspirin: The Remarkable Story of a Wonder Drug*. 2004, New York: Bloomsbury USA
35. Sneader, W., *The discovery of aspirin: a reappraisal.* British Medical Journal, 2000. **321**: p. 1591–1594 (doi: 10.1136/bmj.321.7276.1591)
36. DeVita, V. and E. Chu, *A history of cancer chemotherapy.* Cancer Research, 2008. **68**: p. 8643–8653 (https://doi.org/10.1158/0008-5472.CAN-07-6611)
37. Brogger Chistensen, S., *Drugs that changed society: History and current status of the early antibiotics: salvarsan, sulfonamides, and beta-lactams.* Molecules, 2021. **26**: p. article 6057 (doi: 10.3390/molecules26196057)
38. Dowd, K., *The San Quentin prison doctor who performed over 10,000 human experiments*, in *SFGATE*. 2 May 2022 (https://www.sfgate.com/sfhistory/article/leo-stanley-gland-rejuvenation-surgery-14298920.php).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

39. Lax, E., *The Mold in Dr. Florey's Coat: The Story of the Penicillin Miracle*. 2005, New York: Holt Paperbacks
40. Nicolaou, K.C. and S. Rigol, *A brief history of antibiotics and select advances in their synthesis.* The Journal of Antibiotics, 2018. **17**: p. 153–184 (DOI: 10.1038/ja.2017.62)
41. Einhorn, J., *Nitrogen mustard: the origin of chemotherapy for cancer.* International Journal of Radiation Oncology, Biology, and Physics, 1985. **11**: p. 1375–1378 (https://doi.org/10.1016/0360-3016(85)90254-8)
42. Goodman, L.S., et al., *Nitrogen mustard therapy: use of methyl-bis(beta-chloroethyl)amine hydrochloride and tris(beta-chloroethyl)amine hydrochloride for Hodgkin's Disease, lymphosarcoma, leukemia and certain allied and miscellaneous disorders.* Journal of the American Medical Association, 1946. **132**: p. 126–132 (doi:10.1001/jama.1946.02870380008004)
43. Schwarz, J., *The Right Chemistry: The History of Tylenol*, in *Montreal Gazette*. 2022 (https://montrealgazette.com/opinion/columnists/the-right-chemistry-the-history-of-tylenol).
44. Yanez, M., et al., *History and therapeutic use of MAO-A inhibitors: a historical perspective of mao-a inhibitors as antidepressant drug.* Current Topics in Medicinal Chemistry, 2012. **12**: p. 2275–2282 (DOI: 10.2174/156802612805220011)
45. Braslow, J.T. and S.R. Marder, *History of psychopharmacology.* Annual Review of Psychology, 2019. **15**: p. 25–50 (https://doi.org/10.1146/annurev-clinpsy-050718-095514)
46. Corcoran, A.C., R.D. Taylor, and I.H. Page, *Lithium poisoning from the use of salt substitutes.* Journal of the American Medical Association, 1949. **139**: p. 685–688 (doi:10.1001/jama.1949.02900280001001)
47. Roberts, W.C., *Facts and ideas from anywhere.* Baylor University Medical Center Proceedings, 2015. **28**: p. 421–432 (doi: 10.1080/08998280.2009.11928563)
48. Frishman, W.H., *A historical perspective on the development of alpha-adrenergic blockers.* Journal of Clinical Hypertension, 2007. **9 Suppl 3**: p. 19–27 (doi: 10.1111/j.1524-6175.2007.06633.x)
49. Beyer Jr, K.H., *Discovery of the thiazides: where biology and chemistry meet.* Perspectives in Biology and Medicine, 1977. **20**: p. 410–420 (doi: 10.1353/pbm.1977.0048)
50. Ahlquist, R.P., *A study of the adrenotropic receptors.* American Journal of Physiology, 1948. **153**: p. 586–600 (doi:10.1152/ajplegacy.1948.153.3.586)
51. Godfraind, T., *Discovery and development of calcium channel blockers.* Frontiers in Pharmacology, 2017. **8**: p. article 286 (doi: 10.3389/fphar.2017.00286)
52. Rehman, W., L.M. Arfons, and H.M. Lazarus, *The rise, fall and subsequent triumph of thalidomide: lessons learned in drug development.* Therapeutic Advances in Hematology, 2011. **2**: p. 291–308 (doi: 10.1177/2040620711413165)
53. Greene, J.A. and S.H. Podolsky, *Reform, regulation, and pharmaceuticals - The Kefauver-Harris Amendments at 50.* New England Journal of Medicine, 2014. **367**: p. 1481–1483 (doi: 10.1056/NEJMp1210007)
54. Wikipedia. *History of lysergic acid diethylamide*. 2023; Available from: https://en.wikipedia.org/wiki/History_of_lysergic_acid_diethylamide.
55. Wani, M.C. and S.B. Horwitz, *Nature as a Remarkable Chemist: A Personal Story of the Discovery and Development of Taxolr.* Anticancer Drugs, 2014. **25**: p. 482–487 (doi: 10.1097/CAD.0000000000000063)
56. Gallego-Jara, J., et al., *A compressive review about Taxolr: history and future challenges.* Molecules, 2020. **25**: p. Article 5986 (doi: 10.3390/molecules25245986).
57. Falzone, L., S. Salomone, and M. Libra, *Evolution of cancer pharmacological treatments at the turn of the third millennium.* Frontiers in Pharmacology, 2018. **9**: p. article 1300 (https://doi.org/10.3389/fphar.2018.01300)
58. Smith Hughes, S., *Genentech: The Beginnings of Biotech*. 2013, Chicago: University of Chicago Press
59. Liu, J.K.H., *The history of monoclonal antibody development - Progress, remaining challenges and future innovations.* Annals of Medical Surgery, 2014. **3**: p. 113–116 (doi: 10.1016/j.amsu.2014.09.001)
60. Lyu, X., et al., *The global landscape of approved antibody therapies.* Antibody Therapeutics, 2022. **5**: p. 233–257 (doi: 10.1093/abt/tbac021)
61. GlobeNewsWire, *Monoclonal Antibodies Market is expected to reach USD 612.2 Billion by 2032, growing at a CAGR of 12.3% from 2022 to 2032*. 2023 (https://www.globenewswire.com/news-release/2023/09/06/2738625/0/en/Monoclonal-Antibodies-Market-is-expected-to-reach-USD-612-2-Billion-by-2032-growing-at-a-CAGR-

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

of-12-3-from-2022-to-2032.html).
62. De Clerq, E., *The history of antiretrovirals: Key discoveries over the past 25 years.* Reviews in Medical Virology, 2009. **19**: p. 287–299 (DOI: 10.1002/rmv.624)
63. Berraondo, P., et al., *Cytokines in clinical cancer immunotherapy.* British Journal of Cancer, 2018. **120**: p. 6–15 (doi: 10.1038/s41416-018-0328-y)
64. Matyori, A., et al., *Statins utilization trends and expenditures in the U.S. before and after the implementation of the 2013 ACC/AHA guidelines.* Saudi Pharmacy Journal, 2023. **31**: p. 795–800 (doi: 10.1016/j.jsps.2023.04.002)
65. Gyawali, B. and V. Prasad, *Me-too drugs with limited benefits - the tale of regorafenib for HCC.* Nature Review Clinical Oncology, 2017. **14**: p. 653–654 (DOI: 10.1038/nrclinonc.2017.100)
66. Hughes, A.R., et al., *Genetic association studies to detect adverse drug reactions: abacavir hypersensitivity as an example.* Pharmacogenomics, 2009. **10**: p. 225–233 (https://doi.org/10.2217/14622416.10.2.225)
67. Aronson, J.K. and A.R. Green, *Me-too pharmaceutical products: history, definitions, examples, and relevance to drug shortages and essential medicines lists.* British Journal of Clinical Pharmacology, 2020. **86**: p. 2114–2122 (doi: 10.1111/bcp.14327)
68. National Institute for Health Care Management, *Changing patterns of pharmaceutical innovation*. 2002 (https://nihcm.org/pdf/innovations.pdf).
69. Iqbal, N. and N. Iqbal, *Imatinib: a breakthrough of targeted therapy in cancer.* Chemotherapy Research Practice, 2014: p. article 357027 (doi: 10.1155/2014/357027)
70. American Cancer Society, *Key Statistics for Chronic Myeloid Leukemia*. 2024 (https://www.cancer.org/cancer/types/chronic-myeloid-leukemia/about/statistics.html).
71. Druker, B.J., et al., *Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia.* New England Journal of Medicine, 2001. **344**: p. 1031–1037 (doi: 10.1056/NEJM200104053441401)
72. Paul, M.K. and A.K. Mukhopadhyay, *Tyrosine kinase - role and significance in Cancer.* International Journal of Medical Science, 2004. **1**: p. 101–115 (doi: 10.7150/ijms.1.101)
73. Maadi, H., et al., *Trastuzumab mechanism of action: 20 years of research to unravel a dilemma.* Cancers (Basel), 2021. **13**: p. article 3540 (doi: 10.3390/cancers13143540)
74. Pelechas, E., P.V. Voulgari, and A.A. Drosos, *Preclinical discovery and development of adalimumab for the treatment of rheumatoid arthritis.* Expert Opinion on Drug Discovery, 2021. **16**: p. 227–234 (https://doi.org/10.1080/17460441.2021.1846516)
75. Fu, Z., et al., *Antibody drug conjugate: the "biological missile" for targeted cancer therapy.* Signal Transduction and Targeted Therapy, 2022. **7**: p. article 93 (doi: 10.1038/s41392-022-00947-7)
76. Korman, A.J., K.S. Peggs, and J.P. Allison, *Checkpoint blockade in cancer immunotherapy.* Advances in Immunology, 2006. **90**: p. 297–339 (https://doi.org/10.1016/S0065-2776(06)90008-X)
77. Dobosz, P. and T. Dzieciatkowski, *The intriguing history of cancer immunotherapy.* Frontiers in Immunology, 2019. **10**: p. 16 Dec 2019 (https://doi.org/10.3389/fimmu.2019.02965)
78. Leach, D.R., M.F. Krummel, and J.P. Allison, *Enhancement of antitumor immunity by CTLA-4 blockade.* Science, 1996. **271**: p. 1734–1736 (DOI: 10.1126/science.271.5256.1734)
79. Shalhout, S.Z., et al., *Therapy with oncolytic viruses: progress and challenges.* Nature Reviews Clinical Oncology, 2023. **20**: p. 160–177 (doi: 10.1038/s41571-022-00719-w)
80. Tay, B.Q., et al., *Evolution of cancer vaccines- challenges, achievements, and future directions.* Vaccines (Basel), 2021. **2021 (9)**: p. 535 (doi: 10.3390/vaccines9050535)
81. Wu, S.S., et al., *Reviving an R&D pipeline: a step change in the Phase II success rate.* Drug Discovery Today, 2020. **26**: p. 308–314 (https://doi.org/10.1016/j.drudis.2020.10.019)
82. IQVIA Institute, *Drug Expenditure Dynamics 1995-2020: Understanding medicine spending in context*. 2021 (https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/drug-expenditure-dynamics).
83. Belluck, P., *Congressional inquiry into Alzheimer's drug faults Its maker and F.D.A*, in *New York Times*. 2022 (https://www.nytimes.com/2022/12/29/health/alzheimers-drug-aduhelm-biogen.html).
84. Dunn, B., P. Stein, and P. Cavazzoni, *Approval of aducanumab for Alzheimer Disease-the FDA's perspective.* JAMA Internal Medicine, 2021. **181**: p. 1276–1278 (doi: 10.1001/jamainternmed.2021.4607)
85. Pushpakom, S., et al., *Drug repurposing: progress, challenges and recommendations.* Nature Review Drug Discovery, 2018. **18**: p. 41–58 (https://doi.org/10.1038/nrd.2018.168)
86. CDC Public Health Library, A. Eckert, and A. Higgins. *Ultrastructural morphology of*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

coronaviruses. 2020; Available from: https://phil.cdc.gov/Details.aspx?pid=23311.
87. Heath, D., *Longshot: The Inside Story of the Race for a COVID-19 Vaccine*. 2022, New York: Center Street
88. Heath, D. and G. Garcia-Roberts, *Race to the vaccine*, in *USA Today*. 1–26–2021, USA Today (https://www.usatoday.com/in-depth/news/investigations/2021/01/26/moderna-covid-vaccine-science-fast/6555783002/).
89. Choules, M.P., et al., *Prospective approaches to gene therapy computational modeling - spotlight on viral gene therapy.* Journal of Pharmacokinetics and Pharmacodynamics, 2023. **51**: p. 399–416 (DOI: 10.1007/s10928-023-09889-1)
90. Food and Drug Administration, *FDA at a GLANCE - from the Office of the Commissioner*. 2024 (https://www.fda.gov/media/175664/download).
91. Food and Drug Administration, *Guidance for Industry: Formal Meetings Between the FDA and Sponsors or Applicants of PDUFA Products*. 2023 (https://www.fda.gov/media/172311).
92. Food and Drug Administration. *IND Applications for Clinical Investigations: Regulatory and Administrative Components*. 2022; Available from: https://www.fda.gov/drugs/investigational-new-drug-ind-application/ind-applications-clinical-investigations-regulatory-and-administrative-components.
93. Food and Drug Administration, *Guidance for Clinical Investigators, Sponsors, and IRBs: Investigational New Drug Applications (INDs) - Determining Whether Human Research Studies Can Be Conducted Without an IND*. 2013 (https://www.fda.gov/media/79386/download).
94. Food and Drug Administration, *Content and Format of Investigational New Drug Applications (INDs) for Phase 1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-derived Products*. 1995 (https://www.fda.gov/media/71203/download).
95. Food and Drug Administration. *Federal Food, Drug, and Cosmetic Act (FD&C Act): FD&C Act Reference Information, United States Code, Title 21*. 2018.
96. Camacho, L.H., et al., *Biosimilars 101: considerations for U.S. oncologists in clinical practice.* Cancer Medicine, 2014. **3**: p. article 889 (doi: 10.1002/cam4.258)
97. Liu, T.T., A.S. Kesselheim, and E.R. Scheffer, *Clinical benefit and regulatory outcomes of cancer drugs receiving accelerated approval.* Journal of the American Medical Association, 2024. **331**: p. 1471–1479 (doi:10.1001/jama.2024.2396)
98. Rohrer, M., *Breakthrough Designation Explained*. 2013, Genentech (https://www.gene.com/stories/breakthrough-designation-explained).
99. Tribble, S.J., *High prices For orphan drugs strain families and insurers*, in *National Public Radio*. 1–17–2017 (https://www.npr.org/sections/health-shots/2017/01/17/509507035/high-prices-for-orphan-drugs-strain-families-and-insurers).
100. Kwon, J.M., et al., *An expanded access program of risdiplam for patients with Type 1 or 2 spinal muscular atrophy.* Annals of Clinical Translational Neurology, 2022. **9**: p. 810–818 (doi: 10.1002/acn3.51560)
101. Specter, M., *How ACT UP! Changed America*, in *The New Yorker*. 2021 (https://www.newyorker.com/magazine/2021/06/14/how-act-up-changed-america).
102. Alvaro, D., *From Patients to Partners: The Rising Power of Advocacy in Drug Development*, in *Pharma's Almanac*. 2015 (https://www.pharmasalmanac.com/articles/from-patients-to-partners-the-rising-power-of-advocacy-in-drug-development).
103. Anonymous, *Notes on the Package Insert.* Journal of the American Medical Association, 1969. **207**: p. 1335–1338 (doi:10.1001/jama.1969.03150200101012).
104. Kercik, L. and D.M. Siegel, *Clinical and legal considerations in pharmaceutical compounding.* Journal of Clinical Anesthesiology and Dermatology, 2023. **16(Suppl 8)**: p. S23–S28 (PMCID: PMC10453396).
105. Scola, P.M., et al., *The discovery of asunaprevir (BMS-650032), an orally efficacious NS3 protease inhibitor for the treatment of hepatitis C virus infection.* Journal of Medicinal Chemistry, 2014. **13**: p. 1730–1752 (DOI: 10.1021/jm500297k)
106. Hochaus, A. and H. Kantarjian, *The development of dasatinib as a treatment for chronic myeloid leukemia (CML): from initial studies to application in newly diagnosed patients.* Journal of Cancer Research and Clinical Oncology, 2013. **139**: p. 1971–1984 (DOI: 10.1007/s00432-013-1488-z)
107. World Health Organization, *New INN monoclonal antibody (mAb) nomenclature scheme*. 2021 (https://cdn.who.int/media/docs/default-source/international-nonproprietary-names-(inn)/new_mab_-nomenclature-_2021.pdf).
108. Globus, N.J., *Alphabet Soup: The Story Behind Biosimilar Nonproprietary Name Suffixes*. 2020, Center for Biosimilars (https://www.centerforbiosimilars.com/view/alphabet-soup-

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

the-story-behind-biosimilar-nonproprietary-name-suffixes).
109. Lovelace, B., *'Gaming' of U.S. patent system is keeping drug prices sky high, report says*, in *NBC News*. 2022 (https://www.nbcnews.com/health/health-news/gaming-us-patent-system-keeping-drug-prices-sky-high-report-says-rcna47507).
110. Horrow, C., et al., *Patent portfolios protecting 10 top-selling prescription drugs.* JAMA Internal Medicine, 2024. **184**: p. 810–817 (doi:10.1001/jamainternmed.2024.0836)
111. Downing, N.S., et al., *How Abbott's fenofibrate franchise avoided generic competition.* Archives of Internal Medicine, 2012. **172**: p. 724–730 (doi: 10.1001/archinternmed.2012.187)
112. Boehm, G., et al., *Development of the generic drug industry in the US after the Hatch-Waxman Act of 1984.* Acta Pharmaceutica Sinica B, 2013. **3**: p. 297–311 (https://doi.org/10.1016/j.apsb.2013.07.004)
113. Leo, C.P., et al., *FDA and EMA approvals of new breast cancer drugs-A comparative regulatory analysis.* Cancers (Basel), 2020. **12**: p. article 437 (doi: 10.3390/cancers12020437)
114. Lipinski, C.A., *Lead- and drug-like compounds: the rule-of-five revolution.* Drug Discovery Today, 2004. **1**: p. 337–341 (DOI: 10.1016/j.ddtec.2004.11.007)
115. Shargel, L. and A.B.C. Yu, *Applied Biopharmaceutics and Pharmacokinetics, 4th edition*. 1999, New York: McGraw Hill
116. Yanez, J.A., et al., *Flip-flop pharmacokinetics - delivering a reversal of disposition: challenges and opportunities during drug development.* Therapeutic Delivery, 2011. **2**: p. 643–672 (doi: 10.4155/tde.11.19)
117. Saivin, S., et al., *Clinical pharmacokinetics of acamprosate.* Clinical Pharmacokinetics, 1998. **35**: p. 331–345 (https://doi.org/10.2165/00003088-199835050-00001)
118. Lentz, K.A., *Current methods for predicting human food effect.* AAPS Journal, 2008. **10**: p. 282–288 (doi: 10.1208/s12248-008-9025-8)
119. Benet, L.Z. and B.A. Hoerner, *Changes in plasma protein binding have little clinical relevance.* Clinical Pharmacology and Therapeutics, 2002. **71**: p. 115–121 (DOI: 10.1067/mcp.2002.121829)
120. Zanger, U.M. and M. Schwab, *Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation.* Pharmacology & Therapeutics, 2013. **138**: p. 103–141 (https://doi.org/10.1016/j.pharmthera.2012.12.007)
121. Nebert, D.W. and G. Zhang, *Pharmacogenomics*, in *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics*, R.E. Pyeritz, B.R. Korf, and W.W. Grody, Editors. 2023, Academic Press: New York. p. 445–486
122. McGourty, J.C., et al., *Metoprolol metabolism and debrisoquine oxidation polymorphism - population and family studies.* British Journal of Clinical Pharmacology, 1985. **20**: p. 555–566 (DOI: 10.1111/j.1365-2125.1985.tb05112.x)
123. Rautio, J., et al., *Prodrugs: design and clinical applications.* Nature Review Drug Discovery, 2008. **7**: p. 255–270 (DOI: 10.1038/nrd2468)
124. Olling, M., et al., *Bioavailability of carbamazepine from four different products and the occurrence of side effects.* Biopharmaceutics & Drug Disposition, 1999. **20**: p. 19–28 (DOI: 10.1002/(sici)1099-081x(199901)20:1<19::aid-bdd152>3.0.co;2-q)
125. Ryman, J.T. and B. Meibohm, *Pharmacokinetics of monoclonal antibodies.* CPT Pharmacometrics & Systems Pharmacology, 2017. **6**: p. 576–588 (doi: 10.1002/psp4.12224)
126. Levy, G., *Pharmacologic target-mediated drug disposition.* Clinical Pharmacology and Therapeutics, 1994. **56**: p. 248–252 (DOI: 10.1038/clpt.1994.134)
127. Pollock, C., et al., *Pure red cell aplasia induced by erythropoiesis-stimulating agents.* Clinical Journal of the American Nephrology Society, 2008. **3**: p. 193–199 (DOI: 10.2215/CJN.02440607)
128. Li, J., et al. *Rituximab exhibits a long half-life based on a population pharmacokinetic analysis in Non-Hodgkin's Lymphoma (NHL) patients*. in *American Society of Hematology (ASH) Annual Meeting*. 2007. (https://doi.org/10.1182/blood.V110.11.2371.2371).
129. Schellekens, H., *Bioequivalence and the immunogenicity of biopharmaceuticals.* Nature Review Drug Discovery, 2002. **1**: p. 457–462 (DOI: 10.1038/nrd818)
130. Peletier, L.A. and J. Gabrielsson, *Dynamics of target-mediated drug disposition: characteristic profiles and parameter identification.* Journal of Pharmacokinetics and Pharmacodynamics, 2012. **39**: p. 429–451 (DOI: 10.1007/s10928-012-9260-6)
131. Qi, T., et al., *Cellular kinetics: a clinical and computational review of CAR-T cell pharmacology.* Advanced Drug Delivery Review, 2022. **188**: p. article 114421 (doi: 10.1016/j.addr.2022.114421)

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

132. Stein, A.M., et al., *Tisagenlecleucel Model-Based Cellular Kinetic Analysis of Chimeric Antigen Receptor-T Cells.* CPT Pharmacometrics & Systems Pharmacology, 2019. **8**: p. 295–295 (doi: 10.1002/psp4.12388)
133. Mathijssen, R.H.J., et al., *Flat-fixed dosing versus body surface area-based dosing of anticancer drugs in adults: does it make a difference?* The Oncologist, 2007. **12**: p. 913–923 (DOI: 10.1634/theoncologist.12-8-913)
134. SyBing, A.B., H. Chen, and D.D. Wang, *Re-evaluation of fixed dosing versus body size-based dosing approaches for large molecule therapeutics in adults.* Clinical Pharmacology and Therapeutics, 2025. **15**: p. ahead of print (doi:10.1002/cpt.70015)
135. Freshwater, T., et al., *Evaluation of dosing strategy for pembrolizumab for oncology indications.* Journal of Immunotherapy in Cancer, 2017. **5**: p. 43 (doi: 10.1186/s40425-017-0242-5)
136. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nature Reviews Drug Discovery, 2006. **5**: p. 993–996 (doi: 10.1038/nrd2199)
137. Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nature Review Drug Discovery, 2002. **1**: p. 727–730 (DOI: 10.1038/nrd892)
138. Okuyama, R., *Chronological analysis of first-in-class drugs approved from 2011 to 2022: Their technological trend and origin.* Pharmaceutics, 2023. **15**: p. article 1794 (doi: 10.3390/pharmaceutics15071794)
139. Chartier, M., et al., *Kinome Render: a stand-alone and web-accessible tool to annotate the human protein kinome tree.* Peer Journal, 2013 (https://doi.org/10.7717/peerj.126)
140. Jalencas, X. and J. Mestres, *On the origins of drug polypharmacology.* Medicinal Chemistry Communications, 2013. **4**: p. 80–87 (https://pubs.rsc.org/en/content/articlelanding/2013/md/c2md20242e).
141. Finan, C., et al., *The druggable genome and support for target identification and validation in drug development.* Science Translational Medicine, 2017. **9**: p. eaag1166 (DOI: 10.1126/scitranslmed.aag1166)
142. Dai, S.X., et al., *Proteome-wide prediction of targets for aspirin: new insight into the molecular mechanism of aspirin.* Peer Journal, 2015. **10**: p. e1791 (doi: 10.7717/peerj.1791)
143. Waser, P.G., *The cholinergic receptor.* Journal of Pharmacy and Pharmacology, 1960. **12**: p. 577–594 (DOI: 10.1111/j.2042-7158.1960.tb12714.x)
144. Rubin, R.P., *A brief history of great discoveries in pharmacology: In celebration of the centennial anniversary of the founding of the American Society of Pharmacology and Experimental Therapeutics.* Pharmacological Reviews, 1999. **59**: p. 289–359 (DOI: 10.1124/pr.107.70102)
145. Ritter, J.M., et al., *Rang and Dale's Pharmacology, 9th ed*. 2020, Edinburgh: Elsevier
146. *Consider a typical small molecule drug having a molecular weight of 500 Da and a concentration of 1 ng/mL. Using Avagadro's number, 6.023E23 molecules/mole, the number of molecules is equal to 1.2E12, or a thousand billion molecules per mL. Even for a large protein drug like a monoclonal antibody, having a molecular weight of 150,000 Da, there are 4E9 molecules per mL*. 2025.
147. Liang, Y., S. Li, and L. Chen, *The physiological role of drug transporters.* Protein Cell, 2015. **6**: p. 334–350 (doi: 10.1007/s13238-015-0148-2)
148. Blaschke, G., et al., *Chromatographic separation of racemic thalidomide and teratogenic activity of its enantiomers (author's translation).* Arzneimittelforschung (Drug Research), 1979. **29**: p. 1640–1642 (https://pubmed.ncbi.nlm.nih.gov/583234/).
149. Ruffolo, R.R., *Important concepts in receptor theory.* Journal of Autonomic Pharmacology, 1982. **2**: p. 277–295 (DOI: 10.1111/j.1474-8673.1982.tb00520.x)
150. Rang, H.P., *The receptor concept: pharmacology's big idea.* British Journal of Pharmacology, 2006. **147**: p. S9–S16 (DOI: 10.1038/sj.bjp.0706457)
151. van Dyke, C., et al., *Oral cocaine: Plasma concentrations and central effects.* Science, 1978. **200**: p. 211–213 (DOI: 10.1126/science.24895)
152. Mansoori, B., et al., *The different mechanisms of cancer drug resistance: a brief review.* Advances in Pharmacy Bulletin, 2017. **7**: p. 339–348 (doi: 10.15171/apb.2017.041)
153. Cella, L.K., E.V. Cauter, and D.A. Schoeller, *Diurnal rhythmicity of human cholesterol synthesis: normal pattern and adaptation to simulated jet lag.* American Journal of Physiology, 1995. **269**: p. E489–E498 (doi: 10.1152/ajpendo.1995.269.3.E489)
154. Wang, R., et al., *Antibody-drug conjugates (ADCs): current and future biopharmaceuticals.* Journal of Hematology and Oncology, 2025. **18**: p. article 51 (https://doi.org/10.1186/s13045-025-01704-3)
155. Wang, L., et al., *Therapeutic peptides: current applications and future directions.* Signal

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Transduction and Targeted Therapy, 2022. **7**: p. article 48 (https://doi.org/10.1038/s41392-022-00904-4)
156. Eder, J., R. Sedrani, and C. Wiesmann, *The discovery of first-in-class drugs: origins and evolution.* Nature Review Drug Discovery, 2014. **13**: p. 577–587 (doi: 10.1038/nrd4336)
157. Sadri, A., *Is target-based drug discovery efficient? Discovery and "off-target" mechanisms of all drugs.* Journal of Medicinal Chemistry, 2023. **66**: p. 12651–12677 (https://doi.org/10.1021/acs.jmedchem.2c01737)
158. Fader, L.D., et al., *Discovery of BI 224436, a noncatalytic site integrase inhibitor (NCINI) of HIV-1.* ACS Medicinal Chemistry Letters, 2014. **5**: p. 422–427 (https://doi.org/10.1021/ml500002n)
159. Rappaport, S.M., *Genetic factors are not the major causes of chronic diseases.* PLoS One, 2016. **11**: p. e0154387 (doi: 10.1371/journal.pone.0154387)
160. Minikel, E.V., et al., *Refining the impact of genetic evidence on clinical success.* Nature, 2024. **629**: p. 624–629 (doi: 10.1038/s41586-024-07316-0)
161. Tuzman, K.T., *How to establish causality via patient observation*. 2024 (https://www.biocentury.com/article/653740).
162. Dunn, M.E., et al., *Agonist antibody to guanylate cyclase receptor NPR1 regulates vascular tone.* Nature, 2024. **633**: p. 654–661 (doi: 10.1038/s41586-024-07903-1)
163. Ahmed, A., et al., *Identification of novel serotonin 2C receptor ligands by sequential virtual screening.* Bioorganic Medicinal Chemistry, 2009. **17**: p. 4559–4568 (doi: 10.1016/j.bmc.2009.05.003)
164. Fin, S.F., et al., *A cyclin-dependent kinase inhibitor, dinaciclib in preclinical treatment models of thyroid cancer.* PLoS One, 2017. **12**: p. e0172315 (doi: 10.1371/journal.pone.0172315)
165. van Norman, G.A., *Limitations of animal studies for predicting toxicity in clinical trials: Is it time to rethink our current approach?* JACC Basic & Translational Science, 2019. **4**: p. 845–854 (doi: 10.1016/j.jacbts.2019.10.008)
166. Olson, H., et al., *Concordance of the toxicity of pharmaceuticals and in animals.* Regulatory Toxicology and Pharmacology, 2000. **32**: p. 56–67 (https://doi.org/10.1006/rtph.2000.1399)
167. Russell, W.M.S. and R.L. Burch, *The Principles of Humane Experimental Technique*. 1959, Wheathampstead (UK): Universities Federation for Animal Welfare (https://doi.org/10.5694/j.1326-5377.1960.tb73127.x).
168. *Paul, R. S.5002 - FDA Modernization Act 2.0*. 2022 (https://www.congress.gov/bill/117th-congress/senate-bill/5002/text).
169. Food and Drug Administration, *Roadmap to Reducing Animal Testing in Preclinical Safety Studies*. 2025 (https://www.fda.gov/media/186092/download?attachment).
170. Qureator, I. *First FDA IND milestone achieved using human vascularized organoid efficacy data*. 10/27/2025; Available from: http://ftp.mass-west.com/sponsored/first-fda-ind-milestone-achieved-using-human-vascularized-organoid-efficacy-data.
171. Mennella, J.A., et al., *The bad taste of medicines: overview of basic research on bitter taste.* Clinical Therapeutics, 2013. **35**: p. 1225–1246 (doi: 10.1016/j.clinthera.2013.06.007)
172. Lipchock, S.V., D.R. Reed, and J.A. Mennella, *Relationship between bitter-taste receptor genotype and solid medication formulation usage among young children: a retrospective analysis.* Clinical Therapeutics, 2012. **34**: p. 728–733 (doi: 10.1016/j.clinthera.2012.02.006.)
173. Krietsch Boerner, L., *NDMA, a contaminant found in multiple drugs, has industry seeking sources and solutions.* Chemical and Engineering News, 4–20–2024. **98**(15) (https://cen.acs.org/pharmaceuticals/pharmaceutical-chemicals/NDMA-contaminant-found-multiple-drugs/98/i15).
174. Cantrell, L., et al., *Stability of active ingredients in long-expired prescription medications.* JAMA Internal Medicine, 2012. **172**: p. 1685–1687 (doi:10.1001/archinternmed.2012.4501)
175. Adolph, E.F., *Quantitative relations in the physiological constitutions of mammals.* Science, 1949. **109**: p. 579–585 (DOI: 10.1126/science.109.2841.579)
176. Boxenbaum, H., *Interspecies scaling, allometry, physiological time, and the ground plan of pharmacokinetics.* Journal of Pharmacokinetics and Pharmacodynamics, 1982. **10**: p. 201–207 (DOI: 10.1007/BF01062336)
177. Boxenbaum, H. and R. Ronfeld, *Interspecies pharmacokinetic scaling and the Dedrick plots.* American Journal of Physiology, 1983. **245**: p. R768–R775 (https://doi.org/10.1152/ajpregu.1983.245.6.R768)
178. Dedrick, R., K.B. Bischoff, and D.S. Zaharko, *Interspecies correlation of plasma*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

concentration history of methotrexate (NSC-740). Cancer Chemotherapy Reports, 1970. **54**: p. 95–101 (https://pubmed.ncbi.nlm.nih.gov/5003581/).
179. West, G.B., J.H. Brown, and B.J. Enquist, *A general model for the origin of allometric scaling laws in biology.* Science, 1997. **276**: p. 122–126 (DOI: 10.1126/science.276.5309.122)
180. Food and Drug Administration, *Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers*. 2005 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/estimating-maximum-safe-starting-dose-initial-clinical-trials-therapeutics-adult-healthy-volunteers).
181. International Conference of, H., *ICH E9 Statistical Principles for Clinical Trials - Scientific guideline*. 1998 (https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials-scientific-guideline).
182. Saber, H., et al., *An FDA oncology analysis of CD3 bispecific constructs and first-in human dose selection* Regulatory Toxicology and Pharmacology, 2017. **90**: p. 144–152 (DOI: 10.1016/j.yrtph.2017.09.001)
183. Mishra, A., S.C. Sarangi, and K. Reeta, *First-in-human dose: current status review for better future perspectives.* European Journal of Clinical Pharmacology, 2020. **76**: p. 1237–1243 (DOI: 10.1007/s00228-020-02924-x)
184. Sharma, V. and J.H. McNeill, *To scale or not to scale: the principles of dose extrapolation.* British Journal of Pharmacology, 2009. **157**: p. 907–921 (doi: 10.1111/j.1476-5381.2009.00267.x)
185. Suntharalingam, G., et al., *Cytokine storm in a Phase 1 Trial of the anti-CD28 monoclonal antibody TGN1412.* New England Journal of Medicine, 2006. **355**: p. 1018–1028 (DOI: 10.1056/NEJMoa063842)
186. Panoskaltsis, N., et al., *Immune reconstitution and clinical recovery following anti-CD28 antibody (TGN1412)-induced cytokine storm.* Cancer Immunology and Immunotherapy, 2021. **70**: p. 1127–1142 (doi: 10.1007/s00262-020-02725-2)
187. Medicines and Healthcare Products Regulatory Agency, *Investigations into adverse incidents during clinical trials of TGN1412 (interim report)*. 2006, Medicines and Healthcare Products Regulatory Agency London (https://www.circare.org/foia5/tgn1412rptfinal_20060525.pdf).
188. Horvath, C.J. and M.N. Milton, *The TeGenero incident and the Duff Report conclusions: a series of unfortunate events or an avoidable event?* Toxicology and Pathology, 2009. **37**: p. 372–383 (doi: 10.1177/0192623309332986)
189. Eastwood, D., et al., *Monoclonal antibody TGN1412 trial failure explained by species differences in CD28 expression on CD4+ effector memory T-cells.* British Journal of Pharmacology, 2010. **161**: p. 512–526 (doi: 10.1111/j.1476-5381.2010.00922.x)
190. Tyrsin, D., et al., *From TGN1412 to TAB08: the return of CD28 superagonist therapy to clinical development for the treatment of rheumatoid arthritis.* Clinical and Experimental Rheumatology, 2016. **34**: p. S45–S48 (https://www.clinexprheumatol.org/abstract.asp?a=10866).
191. Peterschmidtt, M.J., et al., *Safety, tolerability, and pharmacokinetics of eliglustat tartrate (Genz-112638) after single doses, multiple doses, and food in healthy volunteers.* Journal of Clinical Pharmacology, 2013. **51**: p. 695–705 (https://doi.org/10.1177/0091270010372387)
192. Food and Drug Administration, *Guidance for Industry: Assessing the Effects of Food on Drugs in INDs and NDAs - Clinical Pharmacology Considerations*. 2022 (https://www.fda.gov/media/121313/download).
193. Pasqualetti, G., et al., *Healthy volunteers and early phases of clinical experimentation.* European Journal of Clinical Pharmacology, 2010. **66**: p. 647–653 (doi: 10.1007/s00228-010-0827-0)
194. Royal College of Physicians, *Research on healthy volunteers: A report of the Royal College of Physicians.* Journal of the Royal College of Physicians, 1986. **20**: p. 243–257 (https://pmc.ncbi.nlm.nih.gov/articles/PMC5371045/).
195. de las Heras, B., et al., *Healthy volunteers in first-in-human oncology drug development for small molecules.* British Journal of Clinical Pharmacology, 2021(88): p. 1773 (https://doi.org/10.1111/bcp.15092)
196. Kabisch, M., et al., *Randomised controlled trials.* Deutsches Arzteblatt, 2011. **108**: p. 663–668 (doi: 10.3238/arztebl.2011.0663)
197. Fogel, D.B., *Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review.* Contemporary Clinical Trials Community, 2018. **11**: p. 156–164 (doi: 10.1016/j.conctc.2018.08.001)
198. Food and Drug Administration, *22 Case Studies Where Phase 2 and Phase 3 Trials Had*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Divergent Results. 2017 (https://www.fda.gov/media/102332/download).
199. McCoy, C.E., *Understanding the intention-to-treat principle in randomized controlled trials.* Western Journal of Emergency Medicine, 2018. **18**: p. 1075–1078 (doi: 10.5811/westjem.2017.8.35985)
200. Lukina, E., et al., *A phase 2 study of eliglustat tartrate (Genz-112638), an oral substrate reduction therapy for Gaucher disease type 1.* Blood, 2010. **116**: p. 893–899 (doi: 10.1182/blood-2010-03-273151)
201. ClinicalTrials.Gov, *Trends, Charts, and Maps*. 2021 (https://classic.clinicaltrials.gov/ct2/resources/trends).
202. Parexel, *Biopharmaceutical R&D Statistical Sourcebook*, ed. M.P. Mathieu. 2015, Waltham, MA: Parexel. 345
203. Cox, T.M., et al., *Eliglustat compared with imiglucerase in patients with Gaucher's disease type 1 stabilised on enzyme replacement therapy: a phase 3, randomised, open-label, non-inferiority trial.* The Lancet, 2015. **383**: p. 2355–2362 (doi: 10.1016/S0140-6736(14)61841-9)
204. Mole, B., *Trump admin axed 383 active clinical trials, dumping over 74K participants*, in *ARS Technica*. 17 Nov 2025 (https://arstechnica.com/health/2025/11/over-74000-people-were-kicked-out-of-clinical-trials-because-of-trump-cuts/).
205. Tall, A.R., L. Yvan-Charvet, and N. Wang, *The Failure of torcetrapib: was it the molecule or the mechanism?* Arteriosclerosis, Thrombosis, and Vascular Biology, 2007. **27**: p. 257–260 (https://doi.org/10.1161/01.ATV.0000256728.60226.77).
206. Clinical Trials Arena Team. *Pipeline Moves: Sanofi terminates breast cancer trial, approval prospects plummet*. 30 Aug 2022; Available from: https://www.clinicaltrialsarena.com/features/pipeline-moves-sanofi-breast-cancer/.
207. Collins, G., et al., *An analysis of dosing-related postmarketing requirements for novel oncology drugs approved by the US Food and Drug Administration, 2012-2022.* Clinical Cancer Research, 2024. **30**: p. 937–941 (https://doi.org/10.1158/1078-0432.CCR-23-2268)
208. Desai, A., et al. *Predictions of isavuconazonium sulfate dosage in pediatrics aged 6 months - 18 years by physiologically based pharmacokinetic model*. in *ID Week*. San Francisco (doi: 10.1093/ofid/ofy210.1227).
209. Olbe, L., E. Carlsson, and P. Lindberg, *A proton-pump inhibitor expedition: the case histories of omeprazole and esomeprazole.* Nature Review Drug Discovery, 2003. **2**: p. 132–139 (doi: 10.1038/nrd1010)
210. Katz, G., et al., *Three-month randomized trial of fixed-combination brinzolamide, 1%, and brimonidine, 0.2%.* JAMA Opthalmology, 2013. **131**: p. 724–730 (DOI: 10.1001/jamaophthalmol.2013.188)
211. Nguyen, Q.H., et al., *Phase 3 randomized 3-month trial with an ongoing 3-month safety extension of fixed-combination brinzolamide 1%/brimonidine 0.2%.* Journal of Ocular Pharmacology and Therapeutics, 2013. **29**: p. 290–297 (DOI: 10.1089/jop.2012.0235)
212. Capati, V.C. and A.S. Kesselheim, *Drug product life-cycle management as anticompetitive behavior: the case of memantine.* Journal of Managed Care and Specialty Pharmacy, 2016. **22**: p. 339–344 (doi: 10.18553/jmcp.2016.22.4.339)
213. Boulanger, V., et al., *Establishing patient registries for rare diseases: rationale and challenges.* Pharmaceutical Medicine, 2020. **34**: p. 185–190 (doi: 10.1007/s40290-020-00332-1)
214. Food and Drug Administration, *Guidance for Industry: Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products*. 2023 (https://www.fda.gov/media/154449/download).
215. Kolb, S.J., et al., *Natural history of infantile-onset spinal muscular atrophy.* Annals of Neurology, 2017. **82**: p. 883–891 (https://doi.org/10.1002/ana.25101)
216. Harrison, R.K., *Phase II and phase III failures: 2013-2015.* Nature Review Drug Discovery, 2016. **15**: p. 817–818 (https://doi.org/10.1038/nrd.2016.184)
217. Sacks, L.V., H.H. Shamsuddin, and Y.I. Yasinskaya, *Scientific and regulatory reasons for delay and denial of FDA approval of initial applications for new drugs, 2000-2012.* Journal of the American Medical Association, 2014. **311**: p. 378–384 (doi:10.1001/jama.2013.282542)
218. Lurie, P., et al., *Comparison of content of FDA letters not approving applications for new drugs and associated public announcements from sponsors: cross sectional study.* British Medical Journal, 2015. **350**: p. h2758 (doi: https://doi.org/10.1136/bmj.h2758)
219. Waldron, J. and F. Kansteiner, *'Weak' evidence and an 'unpleasant' odor: FDA sheds light on drug refusal process*, in *Fierce Biotech*. 7–11–2025

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

(https://www.fiercepharma.com/pharma/weak-evidence-and-unpleasant-odors-fda-sheds-light-drug-refusal-process).
220. *United States Code of Federal Regulations,*, in *Food and Drug Administration Modernization Act of 1997*. 1997 (https://www.govinfo.gov/content/pkg/PLAW-105publ115/pdf/PLAW-105publ115.pdf).
221. Food and Drug Administration, *Guidance for Industry: Demonstrating Substantial Evidence of Effectiveness With One Adequate and Well-Controlled Clinical Investigation and Confirmatory Evidence*. 2023 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/demonstrating-substantial-evidence-effectiveness-one-adequate-and-well-controlled-clinical).
222. Food and Drug Administration, *Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products*. 1998 (https://www.fda.gov/media/71655/download).
223. Food and Drug Administration, *Guidance for Industry: Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products*. 2019 (https://www.fda.gov/media/133660/download).
224. Food and Drug Administration, *Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. 2009 (https://www.fda.gov/media/77832/download).
225. Lederman, S., et al., *Fezolinetant for treatment of moderate-to-severe vasomotor symptoms associated with menopause (SKYLIGHT 1): a phase 3 randomised controlled study.* The Lancet, 2022. **401**: p. 1091–1102 (DOI: 10.1016/S0140-6736(23)00085-5)
226. Chan, A.W., et al., *SPIRIT 2025 statement: updated guideline for protocols of randomized trials.* Nature Medicine, 2025. **31**: p. 1784–1792 (https://doi.org/10.1038/s41591-025-03668-w)
227. Madabushi, R., et al., *Role of model-informed drug development approaches in the lifecycle of drug development and regulatory decision-making.* Pharmaceutical Research, 2022. **39**: p. 1669–1680 (doi: 10.1007/s11095-022-03288-w)
228. Food and Drug Administration, *Product Development Under the Animal Rule*. 2015 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/product-development-under-animal-rule).
229. Momper, J.D., et al., *Ethical considerations for pediatric placebo-controlled trials: FDA outcomes and perspectives.* Therapeutic Innovations in Regulatory Science, 2021. **55**: p. 282–303 (DOI: 10.1007/s43441-020-00214-3)
230. Food and Drug Administration, *Framework for FDA's Real-World Evidence Program*. 2018 (https://www.fda.gov/media/120060/download?attachment).
231. Food and Drug Administration, *Real-World Evidence*. 2025 (https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence).
232. International Conference for Harmonisation, *Choice of Control Group and Related Issues in Clinical Trials (E10)*. 2000 (https://database.ich.org/sites/default/files/E10_Guideline.pdf).
233. Drugs@FDA: Approved Drugs, *Fotivda*. 2021 (https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=212904).
234. Evans, J.M., et al., *Metformin and reduced risk of cancer in diabetic patients.* British Medical Journal, 2005. **330**: p. article 1304 (doi: https://doi.org/10.1136/bmj.38415.708634.F7)
235. Kasznicki, J., A. Sliwinska, and J. Drzewoski, *Metformin in cancer prevention and therapy.* Annals of Translational Medicine, 2014. **2**: p. article 57 (doi: 10.3978/j.issn.2305-5839.2014.06.01)
236. Yu, O.H.Y. and S. Suissa, *Metformin and cancer: solutions to a real-world evidence failure.* Diabetes Care, 2023. **46**: p. 904–912 (https://doi.org/10.2337/dci22-0047)
237. Yadav, K. and R.J. Lewis, *Immortal time bias in observational studies.* Journal of the American Medical Association, 2021. **325**: p. 686–687 (doi:10.1001/jama.2020.9151)
238. Suissa, S. and L. Azoulay, *Metformin and the risk of cancer: time-related biases in observational studies.* Diabetes Care, 2012. **35**: p. 2665–2673 (DOI: 10.2337/dc12-0788)
239. Cejuela, M., et al., *Metformin and breast cancer: where are we now?* International Journal of Molecular Science, 2022. **23**: p. article 2705 (doi: 10.3390/ijms23052705)
240. Giardina, C., et al., *Adverse drug reactions in hospitalized patients: results of the FORWARD (Facilitation of Reporting in Hospital Ward) Study.* Frontiers in Pharmacology, 2018. **9**: p. article 350 (doi: 10.3389/fphar.2018.00350)
241. Monahan, B.P., et al., *Torsade de pointes occurring in association with terfenadine use.* Journal of the American Medical Association, 1990. **264**: p. 2788–2790 (doi:10.1001/jama.1990.03450210088038)

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

242. Kalisch, L.M., et al., *The prescribing cascade.* Australian Prescriber, 2011. **34**: p. 162–166 (https://www.nps.org.au/assets/bfed344b681dd636-b041584d99e7-ca3f75bcf6b64e70d9dde28acdc1db3afe4ec255b4828f2084ec8c611b0c.pdf).
243. Alomar, M.J., *Factors affecting the development of adverse drug reactions.* Saudi Pharmacy, 2014. **22**: p. 83–94 (DOI: 10.1016/j.jsps.2013.02.003)
244. Catita, M., A. Aguas, and P. Morgado, *Normality in medicine: A critical review.* Philosophy, Ethics, and Humanities in Medicine, 2020. **15**: p. article 3 (https://doi.org/10.1186/s13010-020-00087-2)
245. Koeslag, J.H., *"What is Normal?".* South African Medical Journal, 1993. **83**: p. 47–50 (https://www.ajol.info/index.php/samj/article/view/157816).
246. Johnson, R.A., et al., *Risks of phase I research with healthy participants: A systematic review.* Clinical Trials, 2015(13): p. article 149 (doi: 10.1177/1740774515602868)
247. Food and Drug Administration, *Guidance for Industry: Meta-Analyses of Randomized Controlled Clinical Trials to Evaluate the Safety of Human Drugs or Biological Products*. 2018 (https://www.fda.gov/media/117976/download).
248. Ridker, P.M., et al., *Lipid-reduction variability and antidrug-antibody formation with bococizumab.* New England Journal of Medicine, 2017. **376**: p. 1517–1526 (DOI: 10.1056/NEJMoa1614062)
249. Adams, B., *Pfizer dumps PCSK9 inhibitor bococizumab after finding 'no value' in drug*, in *Fierce Biotech*. 2016 (https://www.fiercebiotech.com/biotech/pfizer-dumps-pcsk9i-inhibitor-bococizumab-after-finding-no-value-med).
250. Kirkaesaether Brun, M., et al., *Clinical consequences of infliximab immunogenicity and the effect of proactive therapeutic drug monitoring: exploratory analyses of the randomised, controlled NOR-DRUM trials.* Lancet Rheumatology, 2024. **6**: p. E226–E236 (DOI: 10.1016/S2665-9913(23)00341-7)
251. Aston, R., et al., *Antibody-mediated enhancement of growth hormone activity: application to animal production.* Proceedings of the New Zealand Society of Animal Production, 1991. **51**: p. 227–234 (https://www.nzsap.org/proceedings/1991/antibody-mediated-enhancement-growth-hormone-activity-application-animal-production).
252. Aston, R., W.B. Cowden, and G.L. Ada, *Antibody-mediated enhancement of hormone activity.* Molecular Immunology, 1989. **26**: p. 435–446 (https://doi.org/10.1016/0161-5890(89)90103-X)
253. Vandivort, T.C., D.B. Horton, and S.B. Johnson, *Regulatory and strategic considerations for addressing immunogenicity and related responses in biopharmaceutical development programs.* Journal of Clinical and Translational Science, 2020. **4**: p. 547–555 (doi: 10.1017/cts.2020.493)
254. van Brummelen, E.M.J., et al., *Antidrug antibody formation in oncology: Clinical relevance and challenges.* The Oncologist, 2016. **21**: p. 1260–1268 (doi: 10.1634/theoncologist.2016-0061)
255. Gunn III, G.R., et al., *From the bench to clinical practice: understanding the challenges and uncertainties in immunogenicity testing for biopharmaceuticals.* Clinical and Experimental Immunology, 2016. **184**: p. 137–146 (doi: 10.1111/cei.12742)
256. Food and Drug Administration, *Guidance for Industry: Immunogenicity Information in Human Prescription Therapeutic Protein and Select Drug Product Labeling - Content and Format*. 2022 (https://www.fda.gov/media/155871/download).
257. Food and Drug Administration, *Guidance for Industry: Benefit-Risk Assessment for New Drug and Biological Products*. 2023 (https://www.fda.gov/media/152544/download).
258. Food and Drug Administration, *Risk Evaluation and Mitigation Strategies | REMS*. 2023 (https://www.fda.gov/drugs/drug-safety-and-availability/risk-evaluation-and-mitigation-strategies-rems).
259. Fiala, O., et al., *Skin rash as useful marker of erlotinib efficacy in NSCLC and its impact on clinical practice.* Neoplasma, 2013. **60**: p. 26–32 (DOI: 10.4149/neo_2013_004)
260. Food and Drug Administration, *Good Review Practice: Clinical Review of Investigational New Drug Applications*. 2013 (https://www.fda.gov/media/87621/download).
261. Thunbo, M.O., et al., *Use of medication in pregnancy on the rise: Study on 1.4 million Danish pregnancies from 1998 to 2018.* Acta Obstetrica et Gynecologica Scandinavica, 2024. **103**: p. 1210–1223 (DOI: 10.1111/aogs.14805)
262. Mitchell, A.A., et al., *Medication use during pregnancy, with particular focus on prescription drugs: 1976-2008.* American Journal of Obstetrics and Gynecology, 2011. **205**: p. 51.e1–51.e8 (https://doi.org/10.1016/j.ajog.2011.02.029)
263. Kazma, J.M., et al., *Anatomical and physiological alterations of pregnancy.* Journal of Pharmacokinetics and Pharmacodynamics, 2020. **47**: p. 271–285 (doi: 10.1007/s10928-

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

020-09677-1)
264. Hebert, M.F., et al., *Effects of pregnancy on CYP3A and p-glycoprotein activities as measured by disposition of midazolam and digoxin: a University of Washington Specialized Center of Research Study.* Clinical Pharmacology and Therapeutics, 2008. **84**: p. 248–253 (doi: 10.1038/clpt.2008.1)
265. Schagen, M.R., et al., *Modelling changes in the pharmacokinetics of tacrolimus during pregnancy after kidney transplantation: A retrospective cohort study.* British Journal of Clinical Pharmacology, 2023. **90**: p. 176–188 (https://doi.org/10.1111/bcp.15886)
266. Atkinson, H.C., E.J. Begg, and B.A. Darlow, *Drugs in human milk: clinical pharmacokinetic considerations.* Clinical Pharmacokinetics, 1988. **14**: p. 217–240 (DOI: 10.2165/00003088-198814040-00003)
267. George, B., et al., *Application of physiologically based pharmacokinetic modeling for sertraline dosing recommendations in pregnancy.* NPJ Systems Biology and Applications, 2020. **6**: p. article 36 (doi: 10.1038/s41540-020-00157-3)
268. Cockroft, D.W. and M.H. Gault, *Prediction of creatinine clearance from serum creatinine.* Nephron, 1973. **16**: p. 31–41 (DOI: 10.1159/000180580)
269. Levey, A.S., et al., *Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate.* Annals of Internal Medicine, 2006. **15**: p. 247–254 (DOI: 10.7326/0003-4819-145-4-200608150-00004)
270. Inker, L.A., et al., *New creatinine- and cystatin C-based equations to estimate GFR without race.* New England Journal of Medicine, 2021. **385**: p. 1737–1749 (DOI: 10.1056/NEJMoa2102953)
271. Grasela, D.M., et al., *Pharmacokinetics of single-dose oral stavudine in subjects with renal impairment and in subjects requiring hemodialysis.* Antimicrobial Agents and Chemotherapy, 2000. **44**: p. 2149–2153 (doi: 10.1128/aac.44.8.2149-2153.2000)
272. Zuo, P., et al., *Population pharmacokinetic modeling and exposure-response analysis for the antibody-drug conjugate enfortumab vedotin in locally advanced or metastatic urothelial carcinoma.* Clinical Pharmacology and Therapeutics, 2024. **116**: p. 1278–1288 (doi: 10.1002/cpt.3383)
273. Schwartz, J.B., *Representative enrolment of older adults in clinical trials: the time is now.* Lancet Healthy Longevity, 2023. **4**: p. e301–e304 (DOI: 10.1016/S2666-7568(23)00088-0)
274. Ljungberg, B. and I. Nilsson-Ehle, *Pharmacokinetics of meropenem and its metabolite in young and healthy men.* Antimicrobial Agents and Chemotherapy, 1992. **36**: p. 1437–1440 (DOI: 10.1128/AAC.36.7.1437)
275. Mangoni, A.A. and S.H.D. Jackson, *Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and applications.* British Journal of Clinical Pharmacology, 2004. **57**: p. 6–14 (DOI: 10.1046/j.1365-2125.2003.02007.x)
276. Pomara, N., et al., *Increased sensitivity of the elderly to the central depressant effects of diazepam.* Journal of Clinical Psychiatry, 1985. **46**: p. 185–187 (https://pubmed.ncbi.nlm.nih.gov/3988718/).
277. Vastal, R.E., A.J.J. Wood, and D.G. Shand, *Reduced beta-adrenergic sensitivity in the elderly.* Clinical Pharmacology and Therapeutics, 1979. **26**: p. 181–186 (doi: 10.1002/cpt1979262181)
278. Batchelor, H.K. and J.F. Marriott, *Paediatric pharmacokinetics: key considerations.* British Journal of Clinical Pharmacology, 2015. **79**: p. 395–404 (doi: 10.1111/bcp.12267)
279. Bartelink, I.K., et al., *Guidelines on paediatric dosing on the basis of developmental physiology and pharmacokinetic considerations.* Clinical Pharmacokinetics, 2006. **45**: p. 1077–1097 (doi: 10.2165/00003088-200645110-00003)
280. Phan, H. and M.C. Nahata, *Introduction to Pediatrics, 2nd edition*, in *Pediatric Pharmacotherapy*, S. Benavides and M.C. Nahata, Editors. 2022, American College of Clinical Pharmacy: Lenexa KS. p. 3–17 (ISBN: 1952291038).
281. Wang, C., et al., *Population pharmacokinetics of paracetamol across the human age-range from (pre)term neonates, infants, children to adults.* Journal of Clinical Pharmacology, 2014. **54**: p. 619–629 (doi: 10.1002/jcph.259)
282. Centers for Disease Control and Prevention, *Adult Obesity Facts*. 2024 (https://www.cdc.gov/obesity/php/data-research/adult-obesity-facts.html).
283. World Health Organization, *Obesity and Overweight*. 2024 (https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight).
284. Ortiz, V.E. and J. Kwo, *Obesity: physiologic changes and implications for preoperative management*. 2015. p. article 97 (doi: 10.1186/s12871-015-0079-8).
285. Gao, X., et al., *Mechanisms of obesity-induced changes in pharmacokinetics of IgG in rats.*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Pharmaceutical Research, 2023. **40**: p. 1223–1238 (doi: 10.1007/s11095-023-03496-y)
286. Olafuyi, O., et al., *Inter-ethnic differences in pharmacokinetics-is there more that unites than divides?* Pharmacology Research Perspectives, 2021. **9**: p. e00890 (doi: 10.1002/prp2.890)
287. Takahashi, H., et al., *Population differences in S-warfarin metabolism between CYP2C9 genotype-matched Caucasian and Japanese patients.* Clinical Pharmacology and Therapeutics, 2003. **73**: p. 253–263 (doi: 10.1067/mcp.2003.26a)
288. Brill, M.J.E., et al., *Impact of obesity on drug metabolism and elimination in adults and children.* Clinical Pharmacokinetics, 2012. **51**: p. 227–304 (doi: 10.2165/11599410-000000000-00000)
289. Chagnac, A., et al., *Obesity-induced glomerular hyperfiltration: its involvement in the pathogenesis of tubular sodium reabsorption.* Nephrology Dialysis and Transplantation, 2008. **23**: p. 3946–3952 (https://doi.org/10.1093/ndt/gfn379)
290. Holt, V.L., et al., *Body mass index, weight, and oral contraceptive failure risk.* Obstetrics and Gynecology, 2005. **105**: p. 46–52 (doi: 10.1097/01.AOG.0000149155.11912.52)
291. Ingrande, J. and H.J.M. Lemmens, *Anesthetic pharmacology and the morbidly obese patient.* Current Anesthesiology Reports, 2013. **32**: p. 10–17 (doi: 10.1007/s40140-012-0002-5)
292. Bauer, L.A., et al., *Influence of weight on aminoglycoside pharmacokinetics in normal weight and morbidly obese patients.* European Journal of Clinical Pharmacology, 1983. **24**: p. 643–647 (doi: 10.1007/BF00542215)
293. Griggs, J.J., M.E.S. Sorbero, and G.H. Lyman, *Undertreatment of obese women receiving breast cancer chemotherapy.* JAMA Internal Medicine, 2005. **165**: p. 1267–1273 (doi:10.1001/archinte.165.11.1267)
294. Lin, L., et al., *Are novel oral oncolytics underdosed in obese patients?* Cancer Chemotherapy and Pharmacology, 2024. **93**: p. 129–136 (doi: 10.1007/s00280-023-04601-z)
295. Griggs, J.J., et al., *Appropriate systemic therapy dosing for obese adult patients With cancer: ASCO Guideline Update.* Journal of Clinical Oncology, 2021. **39**: p. 2037–2048 (https://doi.org/10.1200/JCO.21.00471)
296. Food and Drug Administration and Center for Drug Evaluation and Research, *Multi-Discipline Review of Enfortumab Vedotin*. 2019. p. 236 (https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/761137Orig1s000MultiDisciplineR.pdf).
297. Samuels, S., et al., *Obesity considerations in pediatric drug development, 2016-2021.* Journal of Clinical Pharmacology, 2023. **63 (Suppl 2)**: p. S18–S24 (https://doi.org/10.1002/jcph.2305)
298. Food and Drug Administration, *Guidance for Industry: Drug Interaction Studies*. 2024 (https://www.fda.gov/media/161199/download).
299. Dechanont, S., et al., *Hospital admissions/visits associated with drug-drug interactions: a systematic review and meta-analysis.* Journal of Pharmacoepidemiology and Drug Safety, 2014. **23**: p. 489–497 (doi: 10.1002/pds.3592)
300. Bailey, D.G., et al., *Ethanol enhances the hemodynamic effects of felodipine.* Clinical and Investigative Medicine, 1989. **12**: p. 357–362 (https://pubmed.ncbi.nlm.nih.gov/2612087/).
301. Bailey, D.G., et al., *Interaction of citrus juice with felodipine and nifedipine.* The Lancet, 1992. **337**: p. 268–269 (doi: 10.1016/0140-6736(91)90872-m)
302. Lundahl, J., et al., *Effects of grapefruit juice ingestion-pharmacokinetics and haemodynamics of intravenously and orally administered felodipine in healthy men.* European Journal of Clinical Pharmacology, 1997. **52**: p. 139–145 (doi: 10.1007/s002280050263)
303. Greenblatt, D.J., et al., *Time course of recovery of cytochrome p450 3A function after single doses of grapefruit juice.* Clinical Pharmacology and Therapeutics, 2003. **74**: p. 121–129 (doi: 10.1016/S0009-9236(03)00118-8)
304. Romiti, N., et al., *Effects of grapefruit juice on the multidrug transporter P-glycoprotein in the human proximal tubular cell line HK-2.* Life Sciences, 2004. **76**: p. 293–302 (https://doi.org/10.1016/j.lfs.2004.06.015)
305. Bailey, D.G., et al., *Naringin is a major and selective clinical inhibitor of organic anion-transporting polypeptide 1A2 (OATP1A2) in grapefruit juice.* Clinical Pharmacology and Therapeutics, 2007. **81**: p. 495–502 (https://doi.org/10.1038/sj.clpt.6100104)
306. Ross, S.A., et al., *Variance of common flavonoids by brand of grapefruit juice.* Fitoterapia, 2000. **71**: p. 154–161 (https://doi.org/10.1016/S0367-326X(99)00131-8)
307. Ho, P.C., et al., *Content of CYP3A4 inhibitors, naringin, naringenin and bergapten in grapefruit and grapefruit juice products.* Pharmaceutica Acta Helvetiae, 2000. **74**: p. 379–

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

385 (https://doi.org/10.1016/S0031-6865(99)00062-X)
308. Chen, M., et al., *Food-drug interactions precipitated by fruit juices other than grapefruit juice: An update review.* Journal of Food and Drug Analysis, 2018. **26 (2 Suppl)**: p. S61–S71 (doi: 10.1016/j.jfda.2018.01.009)
309. Markowitz, J.S., et al., *Effect of St John's wort on drug metabolism by induction of cytochrome P450 3A4 enzyme.* Journal of the American Medical Association, 2003. **290**: p. 1500–1504 (doi: 10.1001/jama.290.11.1500)
310. Bartone, G.W., et al., *Drug interaction between St. John's wort and cyclosporine.* Annals of Pharmacotherpy, 2000. **34**: p. 1013–1016 (doi: 10.1345/aph.10088)
311. Jusko, W.J., *Influence of cigarette smoking on drug metabolism in man.* Drug Metabolism Reviews, 1979. **9**: p. 221–236 (doi: 10.3109/03602537908993892)
312. Hamilton, M., et al., *Effects of smoking on the pharmacokinetics of erlotinib.* Clinical Cancer Research, 2006. **12**: p. 2166–2171 (https://doi.org/10.1158/1078-0432.CCR-05-2235)
313. Koch, K.M., et al., *Effects of food on the relative bioavailability of lapatinib in cancer patients.* Journal of Clinical Oncology, 2009. **27**: p. 1191–1196 (doi: 10.1200/JCO.2008.18.3285)
314. Ratain, M.J. and E.E. Cohen, *The value meal: how to save $1,700 per month or more on lapatinib.* Journal of Clinical Oncology, 2007. **25**: p. 3397–3398 (DOI: 10.1200/JCO.2007.12.0758)
315. Koch, K.M., et al., *The value of label recommendations: how to dose lapatinib.* Journal of Clinical Oncology, 2007. **25**: p. 5331–5332 (https://doi.org/10.1200/JCO.2007.13.8644)
316. Messori, A., *Effect of food on lapatinib pharmacokinetics.* Journal of Clinical Oncology, 2007. **25**: p. 5332 (https://doi.org/10.1200/JCO.2007.13.9345)
317. Ratain, M.J. and E.E. Cohen, *In reply to The value of label recommendations: How to dose lapatinib.* Journal of Clinical Oncology, 2007(25): p. 5334–5335 (https://doi.org/10.1200/JCO.2007.13.9212)
318. Singh, B.N. and B.K. Malhotra, *Effects of food on the clinical pharmacokinetics of anticancer agents.* Clinical Pharmacokinetics, 2004. **43**: p. 1127–1156 (doi: 10.2165/00003088-200443150-00005)
319. Amidon, G.L., et al., *A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability.* Pharmaceutical Research, 1995. **12**: p. 413–420 (doi: 10.1023/a:1016212804288)
320. Gu, C.H., et al., *Predicting effect of food on extent of drug absorption based on physicochemical properties.* Pharmaceutical Research, 2007. **24**: p. 1118–1130 (doi: 10.1007/s11095-007-9236-1)
321. Schmitt, C., et al., *Disease-drug-drug interaction involving tocilizumab and simvastatin in patients with rheumatoid arthritis.* Clinical Pharmacology and Therapeutics, 2011. **89**: p. 735–740 (https://doi.org/10.1038/clpt.2011.35)
322. Bieck, P.R. and K.H. Antonin, *Tyramine potentiation during treatment with MAO inhibitors: brofaromine and moclobemide vs irreversible inhibitors.* Journal of Neural Transmission, 1989. **28 Suppl**: p. 21–31 (https://pubmed.ncbi.nlm.nih.gov/2794993/)
323. Stass, H., M.F. Bottcher, and K. Ochmann, *Evaluation of the influence of antacids and H2 antagonists on the absorption of moxifloxacin after oral administration of a 400mg dose to healthy volunteers.* Clinical Pharmacokinetics, 2001. **40 (suppl 1)**: p. 39–48 (https://doi.org/10.2165/00003088-200140001-00006)
324. Sadowski, D.C., *Drug interactions with antacids.* Drug Safety, 1994. **11**: p. 395–407 (https://doi.org/10.2165/00002018-199411060-00002)
325. Skelley, J.W., et al., *The impact of tirzepatide and glucagon-like peptide 1 receptor agonists on oral hormonal contraception.* Journal of the American Pharmacists Association, 2024. **64**: p. 204–211 (https://doi.org/10.1016/j.japh.2023.10.037)
326. Hauge, C., et al., *Effect of oral semaglutide on the pharmacokinetics of thyroxine after dosing of levothyroxine and the influence of co-administered tablets on the pharmacokinetics of oral semaglutide in healthy subjects: an open-label, one-sequence crossover, single-center, multiple-dose, two-part trial.* Expert Opinion in Drug Metabolism and Toxicology, 2021. **17**: p. 1139–1148 (doi: 10.1080/17425255.2021.1955856)
327. Ciappina, G., et al., *Impact of proton pump inhibitor use on progression-free and overall survival in cancer patients undergoing immune checkpoint inhibitor therapy: A systematic review and meta-analysis of recent studies.* Cancers, 2025. **17**: p. article 2228 (DOI: 10.3390/cancers17132228)
328. Nasrin, S., et al., *Cannabinoid metabolites as inhibitors of major hepatic CYP450 enzymes, with implications for cannabis-drug interactions.* Drug Metabolism and Disposition, 2021. **49**: p. 1070–1080 (doi: 10.1124/dmd.121.000442)
329. Golombek, P.M., M., et al., *Conversion of cannabidiol (CBD) into psychotropic cannabinoids*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

including tetrahydrocannabinol (THC): a controversy in the scientific literature. Toxics, 2020. **8**: p. 41 (doi: 10.3390/toxics8020041)
330. Johnson, E.D., et al., *Cannabidiol gummy products: LC-MS/MS assessment of cannabinoid concentrations.* Cannabis and Cannabinoid Research, 2025. **June 2** (doi: 10.1089/can.2024.0167)
331. So, G.C., et al., *A Phase I trial of the pharmacokinetic interaction between cannabidiol and tacrolimus.* Clinical Pharmacology and Therapeutics, 2024. **117**: p. 716–722 (https://doi.org/10.1002/cpt.3504)
332. Geffrey, A.L., et al., *Drug–drug interaction between clobazam and cannabidiol in children with refractory epilepsy.* Epilepsia, 2015. **56**: p. 1246–1251 (https://doi.org/10.1111/epi.13060)
333. Leino, A.D., et al., *Evidence of a clinically significant drug-drug interaction between cannabidiol and tacrolimus.* American Journal of Transplant, 2019. **19**: p. 2944–2948 (doi: 10.1111/ajt.15398)
334. Haas, D.W., et al., *Effects of ritonavir on indinavir pharmacokinetics in cerebrospinal fluid and plasma.* Antimicrobial Agents and Chemotherapy, 2003. **47**: p. 2131–2137 (DOI: 10.1128/AAC.47.7.2131-2137.2003)
335. Frye, R.F., et al., *Validation of the five-drug "Pittsburgh cocktail" approach for assessment of selective regulation of drug-metabolizing enzymes.* Clinical Pharmacology and Therapeutics, 1997. **62**: p. 365–372 (https://doi.org/10.1016/S0009-9236(97)90114-4)
336. Menon-Andersen, D., et al., *Essential pharmacokinetic information for drug dosage decisions: a concise visual presentation in the drug label.* Clinical Pharmacology and Therapeutics, 2011. **90**: p. 471–474 (https://doi.org/10.1038/clpt.2011.149)
337. Food and Drug Administration, *Paving the Way for Personalized Medicine: FDA's Role in a New Era of Medical Product Development*. 2013 (https://www.fdanews.com/ext/resources/files/10/10-28-13-Personalized-Medicine.pdf).
338. Precision Medicine Coalition, *Personalized Medicine at the FDA: The Scope and Significance of Progress in 2023*. 2023 (https://www.personalizedmedicinecoalition.org/Userfiles/PMC-Corporate/file/report.pdf).
339. Thanasan, S., et al., *Diagnostic accuracy of immunohistochemistry for HER2-positive breast cancer.* Asian Pacific Journal of Cancer Prevention, 2023. **24**: p. 4321–4327 (doi: 10.31557/APJCP.2023.24.12.4321)
340. Vogel, C.L., et al., *First-line Herceptin monotherapy in metastatic breast cancer.* Oncology, 2001. **61 Suppl 2**: p. 37–42 (https://doi.org/10.1159/000055400)
341. Bonate, P.L., et al., *Extrapolation of praziquantel pharmacokinetics to a pediatric population: a cautionary tale.* Journal of Pharmacokinetics and Pharmacodynamics, 2018. **45**: p. 747–762 (doi: 10.1007/s10928-018-9601-1)
342. Garrod, A.E., *The incidence of alcaptonuria: a study in chemical individuality.* The Lancet, 1902. **160**: p. 1616–1620 (https://doi.org/10.1016/S0140-6736(01)41972-6)
343. Kalow, W., *Pharmacogenetics: Heredity and the Response to Drugs*. 1962, Philadelphia: W.B. Saunders Company
344. Eichelbaum, M. and B. Evert, *Influence of pharmacogenetics on drug disposition and response.* Clinical and Experimental Pharmacology and Physiology, 1996. **23**: p. 983–985 (https://doi.org/10.1111/j.1440-1681.1996.tb01154.x)
345. Eichelbaum, M., et al., *Defective N-oxidation of sparteine in man: a new pharmacogenetic defect.* European Journal of Clinical Pharmacology, 1979. **16**: p. 183–187 (doi: 10.1007/BF00562059)
346. Spring, M.D., et al., *Determination of cytochrome P450 isoenzyme 2D6 (CYP2D6) genotypes and pharmacogenomic impact on primaquine metabolism in an active-duty US military population.* The Journal of Infectious Diseases, 2019. **220**: p. 1761–1770 (https://doi.org/10.1093/infdis/jiz386)
347. Li, J., et al., *Warfarin pharmacogenomics.* Pharmacy and Therapeutics, 2009. **34**: p. 422–427 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2799123/).
348. Schwarz, U.I., et al., *Genetic determinants of response to warfarin during initial anticoagulation.* New England Journal of Medicine, 2008. **358**: p. 999–1008 (DOI: 10.1056/NEJMoa0708078)
349. Rieder, M.J., et al., *Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose.* New England Journal of Medicine, 2005. **352**: p. 2285–2293 (DOI: 10.1056/NEJMoa044503)
350. Asiimwe, I.G., et al., *Warfarin dosing algorithms: a systematic review.* British Journal of Clinical Pharmacology, 2021. **87**: p. 1717–1729 (doi: 10.1111/bcp.14608)
351. The International Warfarin Pharmacogenetics Consortium, *Estimation of the warfarin dose*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

with clinical and pharmacogenetic data. New England Journal of Medicine, 2009. **360**: p. 753–764 (DOI: 10.1056/NEJMoa0809329)
352. Food and Drug Administration, *Coumadin Label*. 2018 (https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/009218s118lbl.pdf).
353. Lima, J.J., et al., *Impact of genetic polymorphisms of the alpha2-adrenergic receptor on albuterol bronchodilator pharmacodynamics.* Clinical Pharmacology and Therapeutics, 1999. **65**: p. 519–525 (doi: 10.1016/S0009-9236(99)70071-8)
354. Goncalo, M., *Usefulness of cutaneous provocation tests to study drugs responsible for cutaneous adverse drug reactions.* Current Treatment Options in Allergy, 2019. **6**: p. 112–124 (https://link.springer.com/article/10.1007/s40521-019-0198-4).
355. Ting, S. and S. Schug, *The pharmacogenomics of pain management: prospects for personalized medicine.* Journal of Pain Research, 2016. **9**: p. 49–56 (https://doi.org/10.2147/JPR.S55595)
356. Mauriello, A., et al., *Pharmacogenomics of cardiovascular drugs for atherothrombotic, thromboembolic and atherosclerotic Risk.* Genes, 2023. **14**: p. article 2057 (https://doi.org/10.3390/genes14112057)
357. Perez-Ramirez, C., et al., *Impact of single nucleotide polymorphisms on the efficacy and toxicity of EGFR tyrosine kinase inhibitors in advanced non-small cell lung cancer patients.* Mutation Research, 2019. **781**: p. 63–70 (https://www.sciencedirect.com/science/article/abs/pii/S1383574217301114).
358. Mallal, S., et al., *HLA-B*5701 screening for hypersensitivity to abacavir.* New England Journal of Medicine, 2008. **358**: p. 568–579 (DOI: 10.1056/NEJMoa0706135)
359. Stainsby, C.M., et al., *Abacavir hypersensitivity reaction reporting rates during a decade of HLA-B*5701 screening as a risk-mitigation measure.* Pharmacotherapy, 2019. **39**: p. 40–54 (https://doi.org/10.1002/phar.2196)
360. Dervieux, T. and M.V. Bala, *Overview of the pharmacoeconomics of pharmacogenetics.* Pharmacogenomics, 2006. **7**: p. 1175–1184 (https://doi.org/10.2217/14622416.7.8.1175)
361. Kalow, W., B.K. Tang, and L. Endrenyi, *Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research.* Pharmacogenetics, 1998. **8**: p. 283–289 (doi: 10.1097/00008571-199808000-00001)
362. Chanfreau-Coffinier, C., et al., *Projected prevalence of actionable pharmacogenetic variants and Level A drugs prescribed among US Veterans Health Administration Pharmacy Users.* JAMA Network Open, 2019. **2**: p. e195345 (DOI: 10.1001/jamanetworkopen.2019.5345)
363. Riedl, J.M., et al., *Molecular diagnostics tailoring personalized cancer therapy?an oncologist's view.* Virchows Archives, 2024. **484**: p. 169–169 (doi: 10.1007/s00428-023-03702-7)
364. Di Pilla, A., et al., *The impact of tumor boards on breast cancer care: evidence from a systematic literature review and meta-analysis.* International Journal of Environmental Research and Public Health, 2022. **19**: p. 14990 (doi: 10.3390/ijerph192214990)
365. Specchia, M.L., et al., *The impact of tumor board on cancer care: evidence from an umbrella review.* BMC Health Services Research, 2024. **20**: p. article 73 (https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-020-4930-3).
366. Lauffer, M.C., et al., *Possibilities and limitations of antisense oligonucleotide therapies for the treatment of monogenic disorders.* Communications in Medicine, 2024. **4**: p. article 6 (doi: 10.1038/s43856-023-00419-1)
367. Kim, J., et al., *Patient-customized oligonucleotide therapy for a rare genetic disease.* New England Journal of Medicine, 2019. **381**: p. 1644–1652 (DOI: 10.1056/NEJMoa1813279)
368. Liem, E.B., et al., *Increased sensitivity to thermal pain and reduced subcutaneous lidocaine efficacy in redheads.* Anesthesiology, 2006. **102**: p. 509–514 (doi: 10.1097/00000542-200503000-00006)
369. Robinson, K.C., et al., *Reduced MC4R signaling alters nociceptive thresholds associated with red hair.* Science Advances, 2021. **7**: p. eabd1310 (DOI: 10.1126/sciadv.abd1310)
370. Liem, E.B., et al., *Anesthetic requirement is increased in redheads.* Anesthesiology, 2004. **101**: p. 279–283 (doi: 10.1097/00000542-200408000-00006)
371. Pigott, G. and R. Tabura. *State ex rel. Shikada v. Bristol-Myers Squibb Co.* 2024; Available from: https://www.fdli.org/2024/05/state-ex-rel-shikada-v-bristol-myers-squibb-co/.
372. Anonymous. *State v. Bristol-Myers Squibb Co.* 2023; Available from: https://law.justia.com/cases/hawaii/supreme-court/2023/scap-21-0000363.html.
373. CBS News, *Hawaii secures $700 million settlement in lawsuit over blood-thinner Plavix*. 9 May 2025 (https://www.cbsnews.com/news/hawaii-secures-700-million-dollar-settlement-lawsuit-blood-thinner-plavix/).
374. Stempel, J., *Texas sues Bristol Myers, Sanofi for concealing blood thinner Plavix's alleged*

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

ineffectiveness, in *Reuters*. 2025 (https://www.reuters.com/world/bristol-myers-sanofi-sued-by-texas-over-plavix-2025-11-20/).
375. West, L.J. and C.M. Pierce, *Lysergic acid diethylamide: its effects on a male Asiatic elephant.* Science, 1962. **138**: p. 1100–1103 (DOI: 10.1126/science.138.3545.1100)
376. Lu, Z., et al., *Dose finding in single dose studies by allometric scaling*, in *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*, F.J. Hock and M.R. Gralinski, Editors. 2018, Springer International Publishing AG: New York (https://doi.org/10.1007/978-3-319-68864-0_3)
377. Food and Drug Administration, *Guidance for Industry: Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers* 2005 (https://www.fda.gov/media/72309/download).
378. Food and Drug Administration, *Guidance for Industry: Nonclinical Evaluation for Anticancer Pharmaceuticals* 2010 (https://www.fda.gov/media/73161/download).
379. Rynn, C., et al., *A translational strategy employing physiologically based modelling to predict the pharmacological active dose of RO7119929, an oral prodrug of a targeted cancer immunotherapy TLR7 agonist.* Xenobiotica, 2022. **52**: p. 855–867 (doi: 10.1080/00498254.2022.2116368).
380. Yoo, C., et al., *Phase 1 study of RO7119929 (TLR7 agonist prodrug) in patients with advanced primary or metastatic liver cancers (CT096)*, in *American Association of Cancer Research*. 2023: Orlando, FL (https://medically.roche.com/content/dam/pdmahub/restricted/oncology/aacr-2023/AACR-2023-poster-yoo-phase-1-study-of-RO7119929.pdf).
381. Yadav, R., et al., *Translational PK/PD and the first-in-human dose selection of a PD1/IL15: an engineered recombinant targeted cytokine for cancer immunotherapy.* Frontiers in Pharmacology, 2024. **15**(1380000) (doi: 10.3389/fphar.2024.1380000).
382. Phipps, A. and on behalf of the AZ Project Optimus Advisory Group. *Applications of model-based approaches to select the starting dosages for first in human trials*. in *Optimizing Doses for Oncology Drug Products: Quantitative Approaches to Select Doses for Clinical Trials*. 2024. Washington D.C. (https://www.aacr.org/wp-content/uploads/2024/02/Opening-and-Session-1A-slides.pdf).
383. Betts, A., et al., *A translational quantitative systems pharmacology model for CD3 bispecific molecules: application to quantify T cell-mediated tumor cell killing by P-cadherin LP DART®.* AAPS Journal, 2019. **21**: p. 66 (doi: 10.1208/s12248-019-0332-z).
384. Brown, M.N., et al., *Efficacy and safety of AZD7594, an inhaled non-steroidal selective glucocorticoid receptor modulator, in patients with asthma: a phase 2a randomized, double blind, placebo-controlled crossover trial.* Respiratory Research, 2019. **20**: p. 37 (https://doi.org/10.1186/s12931-019-1000-7)
385. Yuan, Y., et al., *Bayesian optimal interval design: A simple and well-performing design for Phase I oncology trials.* Clinical Cancer Research, 2016. **22**: p. 4291–4301 (DOI: 10.1158/1078-0432.CCR-16-0592)
386. Kilickap, S., et al., *Cemiplimab monotherapy for first-line treatment of patients with advanced NSCLC With PD-L1 expression of 50% or higher: five-year outcomes of EMPOWER-Lung 1.* Journal of Thoracic Oncology, 2025. **20**: p. 941–954 (https://doi.org/10.1016/j.jtho.2025.03.033)
387. Tandon, R., et al., *A double-blind, placebo-controlled, randomized withdrawal study of lurasidone for the maintenance of efficacy in patients with schizophrenia.* Journal of Psychopharmacology, 2016. **30**: p. 69–77 (doi: 10.1177/0269881115620460)
388. D'Agostino, R., *The delayed-start study design.* New England Journal of Medicine, 2009. **36**: p. 1304–1306 (DOI: 10.1056/NEJMsm0904209)
389. Gonzalez-Duarte, A., et al., *The efficacy of pregabalin in the treatment of prediabetic neuropathic pain.* Clinical Journal of Pain, 2016. **32**: p. 927–932 (DOI: 10.1097/AJP.0000000000000339)
390. Food and Drug Administration, *Guidance for Industry: Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products*. 2019 (https://www.fda.gov/media/121320/download).
391. Millen, G.C. and C. Yap, *Adaptive trial designs: what is the continual reassessment method?* Archives of Disease in Childhood - Education and Practice, 2019. **106**: p. 175–177 (https://doi.org/10.1136/archdischild-2019-316931)
392. Ananthakrishnan, R., et al., *An overview of the BOIN design and its current extensions for novel early-phase oncology trials.* Contemporary Clinical Trials Communications, 2022. **28**: p. article 100943 (doi: 10.1016/j.conctc.2022.100943)
393. Liu, M., S.J. Wang, and Y. Ji, *The i3+3 design for phase I clinical trials.* Journal of

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Biopharmaceutical Statistics, 2020. **30**: p. 294–304 (DOI: 10.1080/10543406.2019.1636811)
394. Food and Drug Administration, *Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics*. 2019 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry).
395. Thorlund, K., J. Haggstrom, and E. Mills, *Key design considerations for adaptive clinical trials: a primer for clinicians.* British Medical Journal, 2018. **360**: p. k698 (doi: https://doi.org/10.1136/bmj.k698)
396. Abrams, P., et al., *Combination treatment with mirabegron and solifenacin in patients with overactive bladder: efficacy and safety results from a randomised, double-blind, dose-ranging, phase 2 study (Symphony).* European Urology, 2015. **67**: p. 577–588 (doi: 10.1016/j.eururo.2014.02.012)
397. Mauri, L. and R. D'Agostino, *Challenges in the design and interpretation of noninferiority trials.* New England Journal of Medicine, 2017. **377**: p. 1357–1367 (DOI: 10.1056/NEJMra1510063)
398. Sherkow, J.S., P.J. Zettler, and H.T. Greely, *Is it 'gene therapy'?* Journal of Law Bioscience, 2018. **5**(786-793) (doi: 10.1093/jlb/lsy020)
399. International Conference of Harmonisation, *ICH E6(R3) Guideline for Good Clinical Practice*. 2025 (https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf).
400. Gardner, W., C.W. Lidz, and K.C. Hartwig, *Authors' reports about research integrity problems in clinical trials.* Contemporary Clinical Trials, 2005. **26**: p. 244–251 (doi: 10.1016/j.cct.2004.11.013)
401. George, S.L. and M. Buyse, *Data fraud in clinical trials.* Clinical Investigations (London), 2015. **5**: p. 161–173 (doi: 10.4155/cli.14.116)
402. Ross, D.B., *The FDA and the case of Ketek.* New England Journal of Medicine, 2007. **356**: p. 1601–1604 (DOI: 10.1056/NEJMp078032)
403. Subcommittee on Oversight and Investigations of the Committee on Energy and Commerce, House of Representatives, and 110th Congress (2nd Sessions), *Ketek Clinical Study Fraud: What did Aventis Know?* 2008 (https://www.congress.gov/event/110th-congress/house-event/LC9490/text).
404. Hobbs, B.P., et al., *Seamless designs: Current practice and considerations for early-phase drug development in oncology.* Journal of the National Cancer Institute, 2018. **111**: p. 118–128 (https://doi.org/10.1093/jnci/djy196)
405. Bogin, V., *Master protocols: new directions in drug discovery.* Contemporary Clinical Trials Communications, 2020. **18**: p. article 100568 (https://doi.org/10.1016/j.conctc.2020.100568)
406. Food and Drug Administration, *Guidance for Industry: Master Protocols: Efficient Clinical Trial Design Strategies to Expedite Development of Oncology Drugs and Biologics*. 2022 (https://www.fda.gov/media/120721/download).
407. Food and Drug Administration, *Conducting Clinical Trials With Decentralized Elements*. 2024 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/conducting-clinical-trials-decentralized-elements).
408. Food and Drug Administration, *Guidance for Industry: E9 Statistical Principles for Clinical Trials*. 1998 (https://www.fda.gov/media/71336/download).
409. Bell, M.L., et al., *Handling missing data in RCTs; a review of the top medical journals.* BMC Medical Research Methods, 2014. **14**: p. 118 (doi: 10.1186/1471-2288-14-118)
410. Lieberman, J.A., et al., *Effectiveness of antipsychotic drugs in patients with chronic schizophrenia.* New England Journal of Medicine, 2005. **353**: p. 1209–1223 (DOI: 10.1056/NEJMoa051688)
411. Leon, A.C., et al., *Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology.* Biological Psychiatry, 2006. **59**: p. 1001–1005 (DOI:10.1016/j.biopsych.2005.10.020)
412. International Conference for Harmonisation, *ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials*. 2020 (https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials-scientific-guideline).
413. Kahan, B.C., et al., *The estimands framework: a primer on the ICH E9(R1) addendum.* British Medical Journal, 2024. **384**: p. e076316 (https://doi.org/10.1136/bmj-2023-076316)

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

414. Zhou, Y. and V.M. Lauschke, *The genetic landscape of major drug metabolizing cytochrome P450 genes-an updated analysis of population-scale sequencing data.* The Pharmacogenomics Journal, 2022. **22**: p. 284–293 (doi: 10.1038/s41397-022-00288-2)
415. Hopewell, S., et al., *CONSORT 2025 statement: updated guideline for reporting randomized trials.* Nature Medicine, 2025. **31**: p. 1776–1783 (https://doi.org/10.1038/s41591-025-03635-5)
416. Campbell, G., et al., *Bayesian statistics for medical devices: progress since 2010.* Therapeutic Innovations in Regulatory Science, 2023. **57**: p. 453–463 (doi: 10.1007/s43441-022-00495-w)
417. Food and Drug Administration, *Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials*. 2010 (https://www.fda.gov/media/71512/download).
418. Food and Drug Administration, *Draft Guidance for Industry: Use of Bayesian Methodology in Clinical Trials of Drug and Biological Products*. 2026 (https://www.fda.gov/media/190505/download).

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Index

(-)-*trans*-Δ^9-tetrahydrocannabinol, 302
* allele, 327
3+3 design, 354
3Rs, 197
Abacavir, 329
Abbreviated New Drug Applications, 81
Absolute bioavailability, 119
Absorption, 115
  effect of food, 120
Accelerated approval, 62, 82
Accumulation, 140
Accumulation ratio, 143
Acetaminophen, 131, 133, 277
Act Up!, 89
Adaptive design, 353, 354
ADCs, 58
ADRs, 246
Aduhelm, 62
Adverse Drug Reaction, 246
Adverse Events, 246
  classification, 247
  rare, 253
  risk factors, 248
  severity, 251
Advocacy Groups, 89
AEs, 246
Affinity, 170
Agonist, 159
Albumin, 124
ALCOA, 361
Allegra, 288
Allometric scaling, 199, 225, 277
Alpha level, 359
Alpha-1-acid-glycoprotein, 124
Alternative hypothesis, 359
Amphipathic, 126
ANDA. *See* Abbreviated New Drug Application
Animal Rule, 240
Antacids, 300
Antagonist, 159, 174
Antibiotics, 33
Antibody-drug conjugates, 58
Anti-Drug Antibodies, 256
Apparent volume of distribution, 112
Area under the curve, 110
Aspirin, 32, 136
Assurance, 359
AUC, 144, *See* Area under the curve
AUC_{ss}, 143
Augmentin, 307
Autoinduction, 129
Basket design, 362
Bayesian Optimal Interval Design, 355
Bayesian statistics, 367
Behavioral sensitization, 183
Belmont Report, 241
Benefit-Risk analysis, 259
Beta-blockers, 41
Beta-receptors, 329
BiDil, 331
Bioavailability, 119
Bioequivalence, 22, 146
Bioequivalence studies, 349
Biologic License Applications, 79
Biologics, 11, 48, 54
  immunogenicity, 256
  pharmacokinetics, 149
Biomarkers, 213
Biosimilars, 22, 63
Biotechnology, 48
Bispecific antibody, 61
BLA. *See* Complete Response Letter
Black, James, 42
Blinatumomab, 345
Blinding, 358
Block parallel group design, 348
BMI, 279, 281
Body mass index, 279
BOIN. *See* Bayesian Optimal Interval Design
BOIN design, 351
Boxenbaum, Harold, 200
Brand names, 100
Breakthrough therapy, 83
Calcium channel blockers, 41, 42
Cannabidiol, 302
Carbidopa, 307
CAR-T cells, 17
Case Report Forms, 361
Causality, 249
CBD, 302
Cell therapy
  pharmacokinetics, 152
Chemical equivalence, 21
Chemotherapy, 37
Chiral Switch, 227
Chronopharmacodynamics, 184
Chronopharmacokinetics, 184
Circadian rhythm, 185
Clark's Rule, 223
Clearance, 109, 110, 111
  apparent oral clearance, 111
  extraction-limited, 110
  flow-limited, 110
  organ, 110

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Clinical Pharmacology Group, 208
Clinicaltrials.gov, 217, 365
Clusters of Differentiation, 16
C_{max}, 109, 144
Combination device, 15
Combinatorial chemistry, 53
Common Technical Document, 78
Common Toxicity Criteria for Adverse Events, 251
Companion diagnostics, 315
Competitive antagonist, 170
Complete Case analysis, 364
Complete Response, 237
Complete Response Letter, 231
Composite Endpoint, 237
Compound libraries, 54
Compound screening, 191
Compounding, 97
Control group\, 213
Controlled Substances Act, 99
Cosmetics, 3, 4
Coumadin, 324
COVID-19 vaccine, 65
Creatinine clearance, 271, 272
Cresemba, 225
CRF-, 361
CRISPR, 68
Crossover design, 349, 350
$C_{ss,avg}$, 143
CYP1A2, 129, 294
CYP2D6, 321
CYP3A, 266, 290, 294
CYP3A4, 118
Cytochrome P450, 118
Cytokines, 297
Data Safety Monitoring Board, 72
Dear Doctor Letter, 254
Decentralized trials, 363
Declaration of Helsinki, 241
Dedrick plots, 200
Delayed Start design, 352
Demographics, 363
Department of Health and Human Services, 71
Dietary supplements, 5, 9
Diethylstilbestrol, 249
DILI, 134
Diuretics, 41, 43
Dose Cap, 281
Dose-Response, 173, 195
Double-blind, 213, 358
Drug, 1
Drug absorption, 115
Drug discovery, 27
Drug distribution, 121
Drug elimination, 136
Drug expiration, 199
Drug Facts Label, 93
Drug metabolism, 125
  autoinduction, 129
  deamination, 127
  hydrolysis, 127
  Phase I, 125
  Phase II, 125
  polymorphisms, 129
Drug names, 99
Drug Quality and Security Act, 98
Drug repurposing, 64
Drug-Induced Liver Injury, 134
Drugs of abuse, 99
DSMB, 72
EC_{50}, 174
ECG, 221
ED_{50}, 174
Ehrlich, Paul, 33
Elderly, 273
E_{max}, 174
Enantiomer, 166
Endpoints, 237
Enrichment design, 352
Enterohepatic recycling, 140
Enzymes, 170
Epigenomics, 332
Epoetin Alfa, 151
Eprex, 151
Estimands, 366
Ethnicity, 280
Expanded access, 89
Exploratory endpoint, 237, 357
Exposure, 109
Extraction ratio, 110
Factorial design, 355
Fair Package and Labeling Act, 90
Fast Track, 83
FDA Modernization Act of 1997, 235
FDAMA, 235
Federal Food, Drug and Cosmetic Act, 71
Federal Food, Drug, and Cosmetic Act, 1
Fexofenadine, 288
FFDCA, 1, *See* Federal Food Drug and Cosmetic Act
Fixed Dose Combinations, 228
Flat dosing, 154
Fleming, Alexander, 36
Flip-flop pharmacokinetics, 115
Flow-limited capacity, 110
Food and Drug administration
  interaction with, 73
Food and Drug Administration
  access programs, 89
  Advisory Committee, 78
  approval pathways, 81
  designations, 76
  Manual of Policies and Procedures, 77
  meeting types, 74
  organization, 73
  Refuse to File, 78
Food and Drug Modernization Act of 1997, 83
Food effect, 208, 209, 295
Forest Plot, 310, 311
Fosamax, 165
Frequentist statistics, 367
Fruit Juice, 288
Full Prescribing Information, 92
Functional foods, 9
Futility analysis, 354
Gene therapy, 67

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

Gene translocation, 324
Genentech, 47
Generic drugs, 21
Genome-Wide Association Studies, 192
Genomics, 332
Genotype, 321
Gleevec, 56
G-proteins, 162
Grapefruit juice, 288
GWAS, 192
Half-life, 109, 113, 141, 144
Haplotype, 327
Hatch-Waxman Act, 21
HED, 341
Hepatic impairment, 269
HER2, 319
Herceptin, 319
Heroin, 31
Highest Non-Severely Toxic Dose, 201, 343
Hit, 193
Hit to Lead, 191
HIV drugs, 51, 64
HNSTD, 343, *See* Highest Non-Severly Toxic Dose
Hoffman, Albert, 44
Hoffman, Felix, 31
Hormones, 33
Human Equivalent Dose, 201, 341
Human kinome, 158
Humira, 58, 103
Hysteresis, 175, 178
IC_{50}, 178
IHC, 319
Imatinib, 55, 56
Immortal-Time Bias, 245
Immune checkpoint inhibitors, 60
Immunogenicity, 256
Immunohistochemistry, 319
Inclusion/Exclusion Criteria, 263
Incomplete block design, 350
IND. *See* Investigational New Drug Applications
Institutional Review Board, 72, 356
Instructions for Use, 94
Insulin, 34, 48
Intellectual property, 102
Intent-to-Treat, 215, 356, 365, 367
International Conference on Harmonisation, 72, 241
International Union of Pure and Applied Chemistry, 99
Inter-occasion variability, 148
Intrinsic activity, 170
Investigational New Drug Application, 76, 99
Investigator-Initiated Trials, 229
Ion channel
  ligand-gated, 162
  voltage-gated, 161
Isavuconazole, 225
Isoniazid, 134
Itraconazole, 308
Juvenile toxicology studies, 196
Kefauver-Harris amendment, 235
Kelsey, Frances, 43
Kidney, 136
K_m, 144
K_p, 124
Lactation, 265
Langley, John Newport, 160
Last Observation Carried Forward, 365
Lead, 193
Lifecycle Management, 226
Lipidomics, 332
Liver, 138
Local tolerability studies, 196
LOCF, 365
LogP, 108
LSD, 44
Lymphatic system, 149
Lysergic acid diethylamide, 44
MABEL, 345, *See* Minimum Anticipated Biological Effect Level
MAD, 210
Magic bullet, 33
MAO inhibitors, 38, 298, 300
MAR, 364
Maria Sharapova, 142
Master Protocol, 362
Maximum Recommended Safe Dose, 200, 342
Maximum Tolerated Dose, 208, 342, 354
MCAR, 364
MED, 344
MedClueRx, 334
Medication Guide, 94
Meldonium, 142
Meta-analysis, 253
Metabolic Ratio, 132
Metabolomics, 332
Metformin, 245
Michael Jackson, 141
Microphysiological systems, 197
Midazolam, 266
MIDD, 222
Milasen, 335
Minimum Anticipated Biological Effect Level, 201, 345
Minimum Effective Dose, 344
Missing at Random, 364
Missing Completely at Random, 364
Missing data, 363
Monitoring, 360
Monoclonal antibody, 13, 49, 58, 152, 153
  pharmacokinetics, 153
Morphine, 28, 31
mRNA vaccines, 66
MSRD, 342, *See* Maximum Recommended Safe Dose
MTD, 342, 354
Multiple Ascending Dose Study, 210
Multiplicity, 355
Mutation, 322
Naringin, 291
Narrow therapeutic index, 133
National Health and Nutrition Survey, 226
NDA, 77
Nephron, 136
New Approach Methodologies, 197
New Drug Application, 77, 90
NHANES, 226
NMDA, 198

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

No Adverse Event Level, 200
No Observable Adverse Effect Level, 342
NOAEL, 200, 342
Noncompetitive antagonist, 174
Nonignorable missing data, 364
Noninferiority trials, 356
Nonpolar, 126
Normal range, 252
Nucleotide deletions, 322
Nucleotide insertions, 322
Null hypothesis, 359
Nuremberg Code, 241
Nutraceuticals, 9
Obamacare. *See* Patient Protection and Affordable Care Act
Obesity, 278
Octanol-to-water partition coefficient, 108
Off-label, 95
OKT3, 49
Open-label, 213, 358
Opium, 28
Organ-on-a-chip, 197
Orphan Drug Act, 71
Orphan drugs, 88
OTC. *See* Over-the-Counter
Overall Survival, 237
Over-the-Counter, 87
Over-the-counter drugs, 229
Package Insert, 90
PAD, 344
Paracelsus, 24
Parallel group, 348
Partial Response, 237
Participant, 347
Partition coefficient, 124
Patent exclusivity, 103
Patents, 102
Patient Medication Information, 94
Patient Package Inserts, 94
Patient Protection and Affordable Care Act, 22
Patient Reported Outcomes, 237
Patient studies, 211
PBPK, 224, 225, 269
PDUFA, 86
Pediatric Rule, 72
Pediatrics, 276
Penicillin, 36
Permeability, 297
Perpetrator drug, 113
Per-protocol, 356, 367
Per-Protocol, 215
Personalized Medicine, 315, 332
P-glycoprotein, 118, 290, 294
PGP, 290, *See* P-glycoprotein
Pharmacodynamics, 157
  time-dependent, 182
Pharmacogenetics, 320
Pharmacokinetics, 108
  biologics, 149
  nonlinear, 145, 146
  time-invariant, 146
Pharmacologically Active Dose, 344
Pharmacophore, 193
Pharmacovigilance, 249
Phase 1, 205
Phase 2, 212
Phase 3, 218
Phase 4, 222
Phenotype, 321
Phenotypic screens, 52
Philadelphia chromosome, 56
Pittsburgh Cocktail, 308
pKa, 122
Placebo, 236
Plavix, 337
PMCs, 86
Pocock method, 355
Polar, 126
Polymorphism, 129, 321
Polypharmacology, 159
Population pharmacokinetics, 148
Post-marketing Commitment, 86
Post-marketing Requirement, 86
Post-Marketing Requirements, 222
Power, 359, 360
Praziquantel, 167
PRCs, 86
Preclinical development, 195
Predictive biomarker, 352
Pregnancy, 265
Prescribing Information, 90
Prescription Drug User Fee Act, 79, 83, 86
Primary endpoint, 357
Primary Endpoint, 237
Priority Review, 83
Prodrug, 33, 133
Product Highlights, 92
Product Label, 90
Prognostic biomarker, 352
Progression-Free Survival, 237
Progressive Disease, 237
Project Bioshield, 72
Project Orbis, 104
PROs, 237
Protease inhibitors, 306
Protein binding, 124
Proteins, 6, 8, 10
Proteomics, 332
Protocol, 356
Protocol deviation, 357
Public Health Service Act, 72
Pure Food and Drug Act of 1906, 73
Pure Red Cell Aplasia, 151
PV, 249
p-value, 366
QTc interval, 286
Randomization, 359
Randomized withdrawal design, 352
Rapid equilibrium, 174
Rare disease, 88
Real-World Evidence, 239, 243
Receptor theory, 159
Receptors, 33, 165
  dopamine, 164

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

drug-receptor complex, 167
fractional occupancy, 168
kinases, 162
metabotropic, 162
nuclear, 164
occupancy, 167
occupation theory of agonism, 169
rate theory, 169
superfamilies, 160
Recommended Phase 2 Dose, 354
Reformulation, 103, 227
Registry or Registries, 230
Relative bioavailability, 120
REMS, 84
Renal Impairment, 269
Rescue design, 352
Response enrichment, 353
Right to Try, 89
Risk-Benefit analysis, 259
RP2D, 354
Rule of 5, 193
SAE, 247
Safety factor, 201, 342
Sample size, 359
Schizophrenia, 39
Seamless Designs, 362
Secondary Endpoint, 237, 357
Seldane, 285
Sentinel dosing, 207
Serious Adverse Event, 247
Serious Adverse Reaction, 247
Serious Unexpected Serious Adverse Reaction, 358
Severely Toxic Dose, 201
Shikada v. Bristol Meyers Squibb and Sanofi, 337
Signal transduction, 164
Simvastatin, 297
Sinemet, 307
Single Ascending Dose design, 350
Single nucleotide polymorphisms, 322
Single-blind, 358
Slow equilibrium, 174
Smoking, 294
SNPs, 322
Solubility, 116, 297
Somatostatin, 48
Specific Populations, 263
St. John's wort, 292, 294
Stable Disease, 237
Statins, 52, 184
Statistical Analysis Plan, 365
Stavudine, 270
STD10. *See* Severely Toxic Dose
Steady-State, 140
Stereoisomer, 166
Stratified parallel group design, 348
Strong CYP3A Inhibitor, 293
Study objectives, 357
Suicide inhibitor, 173
Sulfa drugs, 35
Superiority trial, 356
Superposition, 140
SUSAR, 247, 358
Suspected Unexpected Serious Adverse Reaction, 247
Tachyphylaxis, 182
Tacrolimus, 305
Tarceva, 294
Tardive dyskinesia, 184
Target identification, 190
Target Product Profile, 194
Target validation, 190
Target-mediated drug disposition, 152
Tartaric acid, 166
Taste receptors, 198
Taxanes, 46
Taxol, 45
Terfenadine, 135, 285
TGN1412, 206
Thalidomide, 43, 167
THC, 302
Three Pillars, 107
TITE-BOIN, 351
T_{max}, 109
TMDD, 152
Tocilizumab, 297
Tolerance, 183
Torsades de Pointes, 286
Toxicology, 195
Transcriptomics, 332
Transporters, 118, 163, 170
Trastuzumab, 319
Triple-blind, 213
Tubular reabsorption, 136
Tubular secretion, 137
Tusko, 341
Tykerb, 295
Tylenol, 38
Type I error, 355, 366
Tyramine, 300
Tyrosine kinase inhibitors, 56
UDP-glucuronyltransferase, 130
Umbrella design, 362
US Pharmacopeia, 96
USAN, 99
Use patents, 103
Victim drug, 113
Vioxx, 191
VKORC1, 326
V_{max}, 144
Volume of distribution, 109, 112, 121
apparent V_d, 112
Volunteers, 347
War on Cancer, 45
Warfarin, 324
Weight-based dosing, 154
WHO-UMC Causality Assessment System, 250
Wiley Act, 71
Within-subject variability, 148

Peter Bonate: Drugs Explained. This work is subject to a Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved—Copyright 2026. Version 01MAR2026

About the Author

Peter Bonate has 30 years of experience in the pharmaceutical industry. He is the Head and Executive Director of the Modeling & Simulation group within the Early Development and Translational Science division at Astellas. He has worked at GlaxoSmithKline, Genzyme, ILEX Oncology, Quintiles, Hoechst Marion Roussel, and Eli Lilly & Co. He received his PhD in Medical Neurobiology from Indiana University. He also received an MS in statistics from the University of Idaho and an MS in Pharmacology from Washington State University. He is a Fellow of the American College of Clinical Pharmacology, the American Association of Pharmaceutical Scientists, and the International Society of Pharmacometrics. He is the Editor-in-Chief of the Journal of Pharmacokinetics and Pharmacodynamics. He has served or is currently serving on the editorial boards for Pharmaceutical Statistics, the Journal of Clinical Pharmacology, Pharmaceutical Research, and the AAPS Journal. He is a frequently invited speaker at conferences and universities, has more than 80 publications in the field of pharmacokinetics and clinical pharmacology, is the co-editor of the four-volume series "Pharmacokinetics in Drug Development," and is the author of the leading textbook "Pharmacokinetic-Pharmacodynamic Modeling and Simulation," currently in its second edition, which has more than 50,000 chapter downloads.

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Version Control

Version	Comments
01MAR2026	First version released. Effective 03/01/2026

Peter Bonate: Drugs Explained. This work is subject to Creative Commons CC-BY-NC-ND license.
Subject to such a license, all rights reserved. Copyright 2026. Version 01MAR2026

Drugs Explained

We rely on medications to treat a wide range of illnesses, from the common cold to serious conditions like cancer. Despite how frequently drugs are used in modern medicine, many people have little understanding of how they are developed or how they work in the body. *Drugs Explained* takes the mystery out of medications. Learn how drugs are developed, how they work inside your body, and what all that fine print on medication labels really means. Finally understand the science behind your medicine cabinet!

Peter L. Bonate, PhD, has 30 years of experience in the pharmaceutical industry, and has contributed to the development of many drugs. He is currently Head of the Modeling and Simulation Group within Early Development and Translational Science at Astellas. He is a Fellow of the American Association of Pharmaceutical Sciences, American College of Clinical Pharmacology, and the International Society of Pharmacometrics. He is Editor-in-Chief of The Journal of Pharmacokinetics and Pharmacodynamics and has published more than 80 papers in pharmacokinetics, clinical pharmacology, and drug development.