Flexible FDA Approval Policies

Taylor C. Corcoran, Fernanda Bravo, Elisa F. Long
UCLA Anderson School of Management, Los Angeles CA 90025
taylor.corcoran.1@anderson.ucla.edu, fernanda.bravo@anderson.ucla.edu, elisa.long@anderson.ucla.edu

Current FDA drug approval standards require pharmaceutical companies to demonstrate substantial evidence of effectiveness derived from adequate and well-controlled clinical trials. Traditionally, the FDA requires clinical trial evidence that is statistically significant at the 5% level, but the agency often uses regulatory discretion when interpreting these standards, particularly for drugs targeting rare or severe conditions. Factors such as the target disease severity, prevalence, and availability of approved drugs are qualitatively considered, yet no quantitative framework currently exists for understanding how such characteristics impact approval standards. We propose a novel queueing network model to analyze the drug approval process, which explicitly incorporates these factors. In addition, we capture *obsolescence*—when newer drugs replace older formulas—through use of a pre-emptive $M/M/K/K$ queue for approved drugs. Given an objective of maximizing expected societal benefit less the costs of committing type I and II errors, we show that the optimal significance level is higher for diseases with lengthy clinical trials, greater attrition rates in the development stage, low intensity of research and development, or low levels of obsolescence amongst drugs on the market. Using publicly available clinical trial data, we estimate model parameters and calculate the optimal significance levels for drugs targeting three diseases: breast cancer, HIV, and hypertension. We find that the traditional significance level of 5% is too stringent for some diseases and too lenient for others. Changing approval standards could spur innovation where needed, bringing life-saving drugs to patients, yet allow poor performing, non-efficacious drugs to exit the market.

Key words: FDA, Drug Approval, Queueing Theory, Healthcare
History:

1. Introduction

Since its establishment in 1906, the U.S. Food and Drug Administration (FDA) has approved over 1,500 novel drugs, with total sales of approved drugs exceeding $310 billion each year (Kinch et al. 2014, IMS Health 2016). Despite undergoing rigorous evaluation, some FDA-approved drugs were subsequently shown to be ineffective or even harmful to patients. In September 2004, the anti-inflammatory drug Vioxx was withdrawn from world markets due to safety concerns over increased risks of heart attack and stroke (Dohrman 2005), after more than 160,000 patients suffered adverse events and 38,000 patients died (DrugWatch 2017). The struggle between providing sick patients with potentially beneficial remedies, while protecting consumers from harmful adverse events plays

a significant role in the FDA's decision-making. In this work, we develop a novel queueing model of the drug approval process, starting from development through evaluation, FDA approval or rejection, and obsolescence or market expiry. Our modeling framework can proffer insights for the FDA's approval policy, by permitting flexible approval standards based on differences in disease severity or prevalence, intensity of research and development (R&D), and the number of alternative treatments available for a target condition.[1]

Under the current FDA policy, pharmaceutical companies must first demonstrate that the candidate drug displays no evidence of adverse effects—known as drug *safety*—and second show improvement in a health outcome related to the target condition—known as drug *efficacy*. While both criteria are important, the FDA's primary mission is to ensure that consumers are not unduly harmed by previously approved or new candidate drugs. Safety and efficacy of candidate drugs are usually established by conducting a series of clinical trials, allowing policymakers to weigh the risk of approving an ineffective drug (known as a *type I error*) against the chance of rejecting an effective drug (*type II error*), using statistical hypothesis testing. Traditionally, the probability of type I error is set to a tolerable level known as the *significance level*, α, and the probability of type II error is adjusted through experimental design such as changing the sample size or decreasing measurement error (Casella and Berger 2002).

Traditional hypothesis testing guidelines for drug approval use a significance level of $\alpha = 5\%$ for all diseases. Applying a constant threshold to all drugs has both benefits and challenges. By not prioritizing one disease over another, and holding all drugs to the same safety and efficacy standards, this policy is impartial. On the other hand, the choice of $\alpha = 5\%$ is arbitrary, and no compelling rationale exists for why this specific value was selected (Sterne and Smith 2001). By controlling only for the probability of a type I error, this policy ignores the asymmetric costs of type I and type II errors across diseases. Rejecting an effective medication for mild pain management, which has many other effective treatment options, for example, is less costly than rejecting an effective drug for Alzheimer's disease, for which few treatments are currently available. A fixed threshold ignores the nuances of clinical trial design (e.g., trial duration, rate of new molecule discovery, rate of attrition), characteristics of the target patient population (e.g., disease prevalence and severity), and the post-approval market (e.g., availability of alternative drugs, market maturity).

In recognition of the limitations of a fixed threshold, the FDA has introduced a multitude of programs and regulations that provide the agency with flexibility when evaluating drugs for

[1] In this paper, we refer to a *drug* as a substance intended to diagnose, cure, treat, or prevent disease; we also use this synonymously with the terms medication, therapy, compound, molecule, or drug candidate. The FDA also regulates the approval of medical devices, which we do not explicitly consider in the present study.

approval. Specifically, the FDA uses regulatory discretion to address some aspects of (i) disease *prevalence*, (ii) disease *severity*, and (iii) the duration of the drug development and approval process.

(i) One regulatory mechanism that considers disease prevalence—the number of individuals afflicted with a disease—is the Orphan Drug Act of 1983. In an attempt to offset the high costs of development, the U.S. Congress established tax credits and market exclusivity rights for companies that create drugs for rare diseases (U.S. Food and Drug Administration 2017a, DiMasi et al. 2003). Nevertheless, substantial variation exists in the number of drugs in development, and rare illnesses are not unique in their lack of viable treatments. According to reports by Pharmaceutical Research and Manufacturers of America (PhRMA), 1.6 million cancer cases were newly diagnosed and more than 800 cancer-related drugs were in development in 2015-2016; in contrast, Alzheimer's disease newly afflicted 476,000 people, yet fewer than 80 experimental compounds were in development (PhRMA 2015b, 2016b). One approach to address this imbalance, which we explore in this paper, is through the FDA's choice of significance level for clinical trial results. By changing the significance level for diseases with few drugs in development, the FDA could make it easier for drugs targeting these diseases to gain approval, thus permitting more drugs to enter the market and ultimately, incentivizing pharmaceutical companies to invest more in these illnesses.

(ii) The FDA also partially considers disease severity, a measure of a disease's impact on both mortality (length of life) and morbidity (quality of life) in a patient population. For example, Lotronex, a drug used to treat irritable bowel syndrome, was voluntarily withdrawn from the market in 2000 after many patients experienced severe adverse reactions. Based on positive patient feedback, however, the FDA reintroduced Lotronex in 2002 with restricted use (U.S. Food and Drug Administration 2016b). The FDA's consideration of disease severity is indicated in Title 21, Subpart E of the Federal Code of Regulations, which states that "patients are generally willing to accept greater risks or side effects from products that treat life-threatening and severely-debilitating illnesses, than they would accept from products that treat less serious illnesses" and that "the benefits of the drug need to be evaluated in light of the severity of the disease being treated" (Code of Federal Regulations 2016).

(iii) The FDA introduced the *Fast Track*, *Accelerated Approval*, *Breakthrough Therapy*, and *Priority Review* programs to address the protracted timeline for drug development and approval, which typically lasts between ten and fifteen years (U.S. Food and Drug Administration 2015). The *Fast Track* program facilitates FDA review of candidate drugs that treat serious conditions and fill an unmet medical need. *Accelerated Approval* allows the FDA to base approval decisions for expedited drugs on surrogate endpoints believed to reasonably predict clinical benefit, but are not themselves measures of clinical benefit. For example, a surrogate endpoint for glaucoma-induced vision loss is intraoccular pressure. The *Breakthrough Therapy* program is designed to

hasten the development and review of drugs that demonstrate a significant clinical improvement over existing therapies. Finally, *Priority Review* is a designation requiring the FDA to take action on a drug application within six months, compared to ten months under standard review. In this paper, we explore a different regulatory policy to address these challenges: vary the FDA's choice of significance level based on the drug development process for each disease.

The decision making process of whether to approve or reject a drug is complex. The FDA considers a variety of factors when judging whether to grant marketing approval, including performing a risk-benefit assessment of the drug under consideration, but these factors are weighed qualitatively, making it difficult to ascertain the relative importance of each factor (U.S. Food and Drug Administration 2017d). By developing a model in which the significance level required for approval explicitly depends on the characteristics of the drug development and approval process, one can discern the quantitative effect of a given factor on the likelihood of approval. Furthermore, the FDA has often been accused of fostering opaque approval policies, and a more objective approach to drug approval could improve transparency of this process.

The contributions of this paper are as follows:

• We develop a queueing theory approach to analyze the drug development and approval process, accounting for characteristics such as the R&D intensity, the rate of attrition, and the duration of clinical trials. In our base model, we represent the drug development process using a series of $M/M/\infty$ queues. We consider an extension using a pre-emptive $M/M/K/K$ queueing model that additionally incorporates obsolescence amongst approved drugs. Our study, to the best of our knowledge, is the first to formulate the drug approval process as a network of queues.

• We determine the optimal significance level that maximizes the expected net benefits from approving and rejecting drugs, which include a per-drug benefit for the correct decision of approving effective drugs as well as per-drug type I and type II error costs for the incorrect decisions of approving ineffective and rejecting effective drugs, respectively. In concordance with Montazerhodjat and Lo (2015), we show that the optimal significance level in both models depends on the costs and benefits of approving and rejecting drugs. Furthermore, we show that when drug obsolescence is considered, the optimal significance level depends on the drug development process and is higher for diseases with lengthy clinical trials, high rates of attrition, and low R&D intensity.

• Using publicly available datasets encompassing all registered clinical trials and FDA drug approvals, we estimate model parameters and determine the optimal significance levels for three high-burden diseases: breast cancer, HIV, and hypertension. We test model robustness and show how the optimal significance level relates to characteristics of the development process as well as the costs and benefits of approving and rejecting drugs. Our numeric results highlight that a one-size-fits-all significance level for drug approval is sub-optimal on a societal level, and future research on this topic should consider both pre- and post-approval drug characteristics.

2. Related Literature

Drug Development and Approval. A large body of literature, particularly in health economics, has examined sources of inefficiency in the current drug development and approval process. Three oft-cited sources are the high costs of developing candidate drugs, the high rates of attrition in the development process, and the lack of transparency in the approval process. The Tufts Centre for the Study of Drug Development (2014) estimates an average total cost of $802 million to $2.5 billion to develop a candidate drug and bring it to market. Between 2003 and 2011, 7.5% of all novel drugs that initiated clinical trials ultimately gained approval, with lack of safety and efficacy accounting for over 60% of all candidate drug failures (Hay et al. 2014). In addition to the high R&D costs faced by pharmaceutical companies, the FDA has been accused of fostering opaque approval policies. Downing et al. (2014) examine the strength of clinical trial evidence supporting the approval of novel drugs from 2005 to 2012. The authors find that, despite the FDA's recommendation that drugs are tested against an active comparator or placebo in two randomized, double-blind trials, over 60% of novel drugs were approved on the basis of a single trial, 10% of trials were not randomized, 20% were not double-blind, and 12% did not employ an active comparator of placebo. While this demonstrates the agency's flexibility in considering a wide range of clinical trial evidence, it also obfuscates the agency's drug approval criteria. These prior studies are descriptive in nature and focus on identifying issues in the current drug approval process and quantifying their financial or health burden. In contrast, our work is more prescriptive and presents an objective modeling framework that could help inform policy decisions.

Surprisingly, minimal research has been conducted on analyzing the FDA's decision making process for drug approval. One recent exception is by Montazerhodjat and Lo (2015), who use Bayesian Decision Analysis to quantify how the FDA's approval policy could depend on the severity and prevalence of a disease. After computing the optimal significance levels for 30 leading causes of premature mortality in the United States, the authors argue that the traditional significance level of $\alpha = 5\%$ is too low for severe illnesses with few treatment options such as pancreatic cancer, and too high for less severe illnesses with relatively effective therapies such as diabetes. They raise important questions regarding the role of disease severity and prevalence in drug approval, but their model ignores some key aspects of the development process. Our work extends this study by accounting for pre-approval characteristics, including R&D intensity, the attrition rate in development, clinical trial duration, and post-approval traits, such as time spent on the market and the availability of alternative treatments. Additionally, we approach the problem using a queueing model and optimization rather than a straightforward cost-minimization framework.

Randomized Controlled Trials. Unlike other consumer products brought to market, one bottleneck in the pharmaceutical drug approval process is the time needed to complete the requisite

sequence of clinical trials. A large body of research focuses on optimal trial design so as to shorten trial duration or minimize the number of volunteers exposed to a potentially unsafe drug. Ahuja and Birge (2016) study the problem of dynamically adjusting randomization probabilities of patients to treatments so that patients are treated as effectively as possible without compromising the ability to learn about treatment efficacy. Bertsimas et al. (2015) employ discrete linear optimization methods to construct treatment groups for small samples to allow for more powerful statistical inference. Small sample clinical trial design is important for ethical reasons, but also logistically, as it is often difficult to recruit a large number of volunteers for rare disease trials. Montazerhodjat and Lo (2015) explicitly incorporate the costs associated with treating patients with a potentially harmful drug and they use expected cost analysis to determine the optimal sample size for a balanced two-arm randomized controlled trial. Yapar et al. (2016) use a Bayesian expected value of information framework to design multi-arm, multi-stage trials allowing drugs to be accepted or rejected earlier than with traditional trials. Other recent studies examine existing clinical trial data to identify which combinations of novel drug compounds are most effective. For example, Bertsimas et al. (2016) use machine learning to predict chemotherapy outcomes in cancer patients and the authors suggest new combinations of drug regimens to test in future trials. Our work differs in that we do not explicitly model clinical trial design, but rather we analyze how disease specifics drive the optimal significance level, assuming a standard balanced two-arm trial design.

New Product Development. The journey of a candidate drug from conception through research and development, testing, regulatory approval, and post-approval market penetration also relates to studies on new product development (NPD). See Krishnan and Ulrich (2001) and Killen et al. (2007) for a comprehensive review of this vast body of literature. NPD is the process of transforming product concepts into commodities that can be sold. As in the FDA setting, both the development process and the market stage of NPD have inspired academic research.

Some studies in the NPD literature examine how time to market is affected by resource-sharing amongst projects in development. Adler et al. (1995) model the product development process as a queueing network, where engineering resources are "servers" and projects are "jobs" that move between stages of development. This framework can identify bottlenecks in development and opportunities to reduce time to market for new products. Our work similarly models the stages of drug development as a sequence of queues. One difference, however, is that we model the pre-approval drug development process using $M/M/\infty$ queues; thus we do not capture resource-sharing imposed by a finite number of servers. Adler et al. (1995) examine product development within a *single* company, whereas we study drug development across *all* pharmaceutical companies, which ignores delays caused by constrained resources within a firm.

Among NPD research on the market stage of development, our work most closely relates to studies that examine how projects targeting a single market compete for profit. Ding and Eliashberg (2002) use dynamic programming to determine the optimal number of projects to pursue to maximize expected profit, when the final products target the same market and compete for revenue. They define the number of projects pursued by a firm as a decision variable, whereas R&D intensity is an estimated parameter in our work. Furthermore, rather than studying competition for revenue, we examine the role of obsolescence amongst drugs targeting the same condition.

Queueing Models. Queueing models have been used extensively to design and analyze systems in industries ranging from manufacturing to health care; comprehensive reviews can be found in Govil and Fu (1999) and Green (2006), respectively. While queueing theory arises naturally in many settings, it is a flexible tool that can be applied to problems where the choice of servers and customers is less obvious. Kaplan (2010), for example, estimates the number of undetected terror plots in a region by modeling potential plots as customers and intelligence officers as servers. Cossin and Schellhorn (2007) analyze credit risk amongst firms using a network queueing model with firms as servers. We also diverge from traditional queueing applications by modeling candidate drugs as "customers" and each stage of the development and post-review process as a queue.

3. Drug Development Overview

The current drug approval process in the U.S. consists of a series of stages, beginning with the discovery of the potential new pharmaceutical compound and ending with the FDA deciding whether to grant marketing approval to the drug. See Figure 1 for a summary.

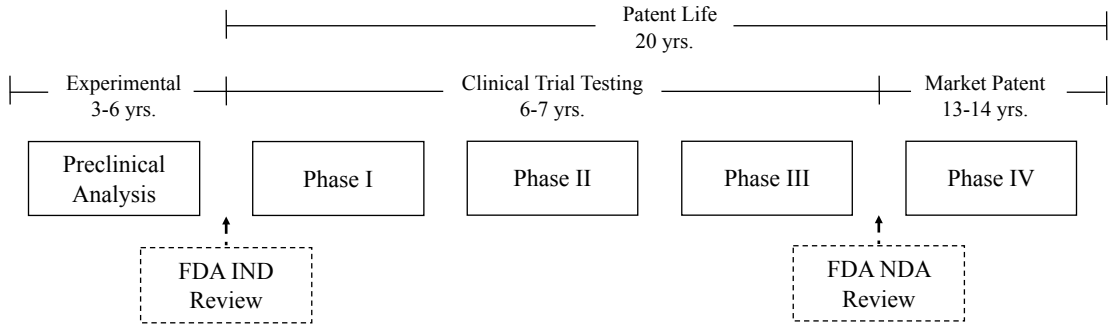
3.1. Role of the FDA

The complex process of creating a new pharmaceutical drug begins with extensive research of the target disease and identification of a novel chemical compound intended to treat the illness. Promising candidates are subjected to preclinical analysis, involving laboratory (*in vitro*) and animal (*in vivo*) testing. In addition to screening for potential safety issues, the purpose of these tests is to study how the candidate drug is eventually processed by the human body (pharmacokinetics) and to determine appropriate dosing levels. On average, drug discovery and preclinical investigation lasts between three and six years (PhRMA 2015a).

If a candidate drug raises no safety concerns during preclinical testing, the sponsoring firm can submit an Investigational New Drug (IND) application to the FDA, presenting preclinical results and plans for performing clinical trial testing. The firm may begin clinical trials within 30 days of filing an IND, provided the FDA does not respond with objections to the proposed testing plan.

Clinical trial testing usually consists of three phases, designed to determine if the candidate drug is both safe and effective in humans, lasting a total of 6 to 7 years, on average (PhRMA 2015a).

Figure 1 The FDA drug development and approval process.



Note. For each drug candidate, a pharmaceutical firm submits two applications to the FDA for review. IND = Investigational New Drug; NDA = New Drug Application.

Phase I trials involve testing the candidate drug on groups of 20 to 80 healthy volunteers, with the goal of observing potential side effects and studying the drug's pharmacokinetics and last 20 months, on average. Provided that the therapy is well-tolerated by healthy volunteers, the drug can advance to Phase II, where it is administered to volunteers who suffer from the target illness. The goal is to establish drug efficacy in sick patients while continuing to monitor side effects. Efficacy is typically demonstrated through the use of randomized controlled trials, in which patients receiving the candidate drug are compared to those treated with a placebo or standard therapy. Phase II trials last an average of 30 months. The final stage of clinical testing, Phase III, entails testing the candidate drug on hundreds or thousands of volunteers with the target disease. As with Phase II, the focus is establishing efficacy in a large patient cohort. Additionally, researchers study how the candidate drug interacts with other medications, how different populations react to the drug, and which dosage levels are practical. The average duration of Phase III trials is 30 months (Tufts Centre for the Study of Drug Development 2014).

At any point during the preclinical or clinical trial stages, a pharmaceutical company may choose to withdraw the drug from development. Typical reasons for halting development include the inability to demonstrate efficacy, concerns over safety, issues with how the drug is metabolized (pharmacokinetics), market competition, and financial considerations (Arrowsmith and Miller 2013). Companies that do not withdraw their product, and whose clinical trial results demonstrate that the candidate drug has a reasonable risk-benefit profile, can submit a New Drug Application (NDA) to the FDA. In this application, the firm must provide results of preclinical and clinical testing as well as a proposal for manufacturing and labeling the medication. The NDA is examined by FDA review teams, who perform a risk-benefit assessment using information from clinical trial testing, including data on demonstrated efficacy and reported adverse events, and decide whether the potential benefits of the medication outweigh its risks. Periodically, companies are asked to

perform additional testing or to revise aspects of their application before being awarded marketing approval (PhRMA 2013, 2015a, U.S. Food and Drug Administration 2014).

Drugs that ultimately gain FDA approval may then be legally marketed in the United States. During their time on the market, some drugs benefit from patents or exclusivity rights. Patents are granted by the U.S. Patent and Trademark Office and typically expire 20 years after a sponsoring firm files a patent application, usually before initiating clinical trial testing, though applications can be submitted at any point during the development process. Patents can be issued for drugs regardless of their approval status. Exclusivity, or exclusive marketing rights, are granted by the FDA, with all new drugs receiving 5 years of exclusivity upon approval. Companies that perform pediatric studies receive an additional 6 months of exclusivity (5.5 years total), and drugs that are initially approved for an orphan indication are granted an additional 2 years (7 years total) (U.S. Food and Drug Administration 2016d). Safety and efficacy of approved drugs continue to be monitored during post-marketing studies (Phase IV), with any adverse events caused by the medication reported to the FDA (PhRMA 2015a, U.S. Food and Drug Administration 2016c). Most drugs do not cause wide-scale adverse events and thus remain on the market as long as the sponsoring firm chooses to continue manufacturing them. In rare cases, drugs with harmful side effects are withdrawn from the market either by the developing firm or the FDA (U.S. Food and Drug Administration 2017b).

3.2. Randomized Controlled Trial Design

When deciding whether to approve a candidate drug, the main criteria used by the FDA are safety and efficacy, both of which are evaluated using clinical trial data. Safety is assessed by considering the number and type of adverse events experienced by trial volunteers. Efficacy is assessed by monitoring one or more clinical endpoints—outcomes that represent direct clinical benefit—associated with the target disease, and determining whether the drug has a statistically significant impact on the endpoint (CDER and CBER 1998). We present a standard statistical framework for modeling drug *efficacy* (Section 3.3) but we do not explicitly model drug *safety* concerns. We make this modeling choice because the number of potentially harmful side effects is large, and these effects are usually unforeseen at the trial's start (Friedman et al. 2015). In contrast, the number of clinical endpoints used to assess efficacy is small, and these endpoints must be specified before initiating the trial.

Randomized controlled trials are the current standard for establishing efficacy of candidate drugs. For simplicity, we assume that all candidate drugs are tested using a two-arm balanced randomized trial, a commonly used design where patients are randomly assigned to a *control* group or a *treatment* group, which are equal in size. Individuals in the treatment arm receive the experimental

regimen; those in the control arm receive standard therapy or a placebo. Before the trial commences, researchers must propose one or more endpoints associated with the target disease that will be monitored throughout the study (Friedman et al. 2015, Jennison and Turnbull 2000). For example, one endpoint for some oncology drugs is five year progression-free survival. In this work, we assume that one quantitative endpoint is monitored, though in reality multiple endpoints can be used, and we assume that higher values of the endpoint are associated with better health outcomes, though a range of desirable values could exist. Measurements of the single quantitative endpoint from patients in both groups are analyzed, and the drug is said to be effective if the response of the treatment group is larger and statistically different from the response of the control group.

3.3. A Statistical Framework for Drug Approval

Consider a two-armed balanced clinical trial with n patients in each arm. Let $X_1, \ldots, X_n$ denote independent observations of a single quantitative endpoint from patients in the treatment group, and let $Y_1, \ldots, Y_n$ denote independent observations from patients in the control group who receive standard therapy. We assume that X_i follows a normal distribution with mean μ_x and variance σ^2, and Y_i follows a normal distribution with mean μ_y and variance σ^2 (Jennison and Turnbull 2000). The assumption of equal variance is made for simplicity and can be easily relaxed.

The quantity $\delta = \mu_x - \mu_y$ represents the treatment effect of the candidate drug. Our analysis focuses on non-inferiority trials, which assumes that the experimental drug has no effect or a positive effect, compared to standard therapy. We perform the following hypothesis test:

$$H_0 : \delta = 0$$
$$H_1 : \delta > 0$$

using the following Wald statistic from the observed data, $X_1, \ldots, X_n, Y_1, \ldots, Y_n$:

$$Z_n = \left(\overline{X} - \overline{Y}\right) \sqrt{I_n}$$

where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ are the sample means, and $I_n = \frac{n}{2\sigma^2}$ is known as the *information* of the sample. It can be shown that Z_n follows a normal distribution with mean $\delta\sqrt{I_n}$ and variance 1. After calculating Z_n, one then computes the p-value associated with the Wald statistic and compares this value to a threshold α. If the p-value is less than α, then H_0 is rejected and the drug is deemed effective. If the p-value is greater than α, the null hypothesis cannot be rejected, and the drug is considered ineffective.

Let the *approval policy* corresponding to significance level α be defined as follows: candidate drugs that complete clinical trials and undergo FDA review are approved if the p-value $< \alpha$, and rejected otherwise. Let p be the *prior* probability that a candidate drug is actually effective (the

alternative hypothesis). Given an approval policy α and prior p, we obtain the following probability expressions:

$$\mathbb{P}(\text{AE}) = [1 - \Phi(\Phi^{-1}(1-\alpha) - \delta\sqrt{I_n})]\, p \qquad \text{[Approving an effective drug]} \tag{1}$$

$$\mathbb{P}(\text{AI}) = \alpha\,(1-p) \qquad \text{[Approving an ineffective drug]} \tag{2}$$

$$\mathbb{P}(\text{RE}) = \Phi(\Phi^{-1}(1-\alpha) - \delta\sqrt{I_n})\, p \qquad \text{[Rejecting an effective drug]} \tag{3}$$

$$\mathbb{P}(\text{RI}) = (1-\alpha)\,(1-p) \qquad \text{[Rejecting an ineffective drug]} \tag{4}$$

Here, Φ and Φ^{-1} are the cumulative distribution function and inverse cumulative distribution function, respectively, of the standard normal.

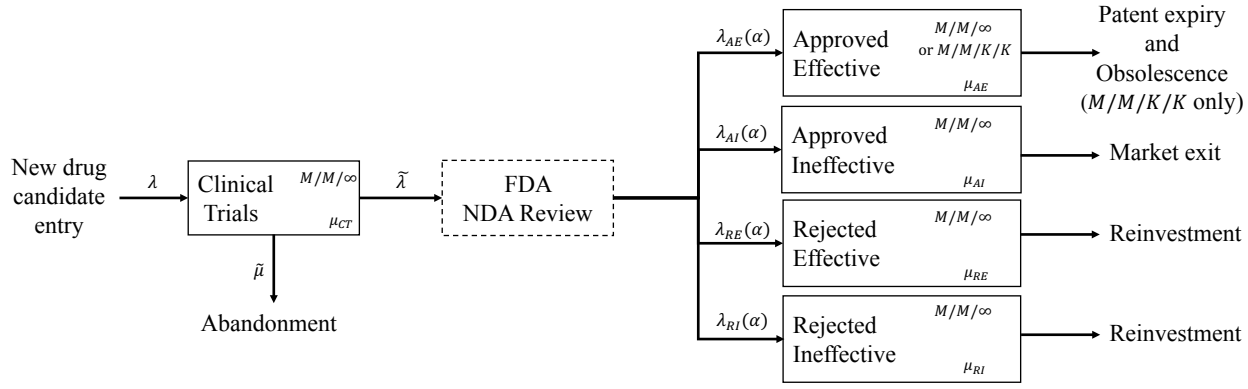
4. A Queueing Framework for the Drug Approval Process

We introduce a queueing network to model the development of a candidate drug from clinical trials to post-approval (Figure 2). A summary of model parameters is provided in Table 1.

4.1. Queueing Network Model

$M/M/\infty$ Model. Candidate drugs begin clinical trials according to a Poisson process with rate λ. Drugs either complete clinical trials successfully, or their sponsoring company decides to halt the trials early. We assume that the time until a drug company abandons the trial is exponentially distributed with rate μ_{AB}, and the time to complete trials is exponentially distributed with rate μ_{CT}. Drugs in clinical trials are considered to be in "service", and after completion of service, drugs either advance to the FDA review stage with probability $\frac{\mu_{CT}}{\mu_{CT}+\mu_{AB}}$ or abandon clinical trials and exit the system with probability $\frac{\mu_{AB}}{\mu_{CT}+\mu_{AB}}$. We denote the net rate at which drugs enter FDA review by $\tilde{\lambda} = \lambda\,\frac{\mu_{CT}}{\mu_{CT}+\mu_{AB}}$ and the net rate at which drugs abandon clinical trials as $\tilde{\mu} = \lambda\,\frac{\mu_{AB}}{\mu_{CT}+\mu_{AB}}$.

Figure 2 **Queueing network representing the drug development and approval process.**



Note. We refer to the case where the approved effective stage is modeled as an $M/M/K/K$ queue as the K-bumping model.

After NDA review, drugs gain FDA approval if the p-value associated with the clinical trial demonstrating efficacy is less than the chosen statistical significance level α, and are denied approval otherwise. For simplicity, we assume that the FDA's decision to approve or reject a drug occurs immediately, though in reality the review process takes between 6 months and 2 years. This delay could be accounted for by modeling the review stage as an $M/M/\infty$ queue, but would not change our results. After undergoing FDA review, drugs probabilistically move to one of four post-review stages. By Burke's theorem, in steady state, the output process of the clinical trials stage is a Poisson process with rate $\tilde{\lambda}$. Thus, the FDA review stage constitutes a thinning of a Poisson process into four separate and independent Poisson processes: approved effective drugs (AE), approved ineffective drugs (AI), rejected effective drugs (RE), and rejected ineffective drugs (RI). Using the probabilities in equations (1)-(4), the arrival rates for each of these Poisson processes are:

$$\lambda_{AE}(\alpha) = \tilde{\lambda}\,[1 - \Phi(\Phi^{-1}(1-\alpha) - \delta\sqrt{I_n})]\,p \tag{5}$$

$$\lambda_{AI}(\alpha) = \tilde{\lambda}\,\alpha\,(1-p) \tag{6}$$

$$\lambda_{RE}(\alpha) = \tilde{\lambda}\,\Phi(\Phi^{-1}(1-\alpha) - \delta\sqrt{I_n})\,p \tag{7}$$

$$\lambda_{RI}(\alpha) = \tilde{\lambda}\,(1-\alpha)\,(1-p) \tag{8}$$

Each post-review stage is modeled as a queue, where "service" represents time on the market (for approved drugs) or time until a pharmaceutical company initiates development of a different drug candidate (for rejected drugs). After completing service, all drugs depart the system. In the simplest case, we model the four post-review stages as $M/M/\infty$ queues with service rates μ_{AE}, μ_{AI}, μ_{RE} and μ_{RI}, respectively. We refer to this queueing network as the $M/M/\infty$ model. Intuitively, the $M/M/\infty$ model views the drug development and approval process as a series of stages where the time required to complete each stage is random and independent across stages and among drugs.

This $M/M/\infty$ formulation assumes exogenous service rates corresponding to time on the market, or time until reinvestment in a new candidate drug. This assumption likely holds for rejected drugs that never enter the market, and for approved ineffective drugs, which are identified by dissatisfied patients. However, approved effective drugs may be at risk of becoming obsolete as new drugs subsequently enter the market. Furthermore, patients typically take only a few drugs to treat a given disease, and drugs within the same pharmaceutical class often function as substitutes.

***M/M/K/K* Model.** To integrate obsolescence and substitution amongst approved effective drugs, we propose a variation of a preemptive queueing model, which we refer to as the K-bumping model. Assume that at most K approved effective drugs are on the market at a time. If fewer than K drugs are in the approved effective stage, then additional approved effective drugs begin service according to a first-come, first-served policy. However, if an effective drug gains approval

but K drugs are already in the approved effective stage, then this new drug replaces (or "bumps") the drug in service the longest. By solving the flow balance equations, we find that the approved effective stage with this "bumping" protocol can be modeled in steady state as an $M/M/K/K$ queue. The bumping of approved effective drugs is designed to capture the phenomenon where older drugs become obsolete as newer therapies gain approval. Drug substitution typically occurs within a pharmaceutical class, assuming patients take at most one drug within a class. Therefore, we associate each effective drug on the market with a drug class and view K as the number of unique drug classes available to treat a particular disease.

Table 1 Summary of key model parameters.

Before FDA review		After FDA review	
σ	Standard deviation of the drug response	K	Number of unique drug classes on the market
δ	Treatment effect of a candidate drug	B_{AE}	Per drug benefit of approving effective drugs
p	Prior probability that candidate drug is effective	C_{AI}	Per drug cost of approving ineffective drugs
n	Clinical trial enrollment	C_{RE}	Per drug cost of rejecting effective drugs
λ	Rate that drugs initiate a clinical trial	$1/\mu_{AE}$	Average duration of approved effective stage
μ_{CT}	Rate that clinical trials are completed	$1/\mu_{AI}$	Average duration of approved ineffective stage
μ_{AB}	Rate that firms abandon clinical trials	$1/\mu_{RE}$	Average duration of rejected effective stage

In both the $M/M/\infty$ and K-bumping models, we assume that the FDA's objective is to choose the statistical significance level α to maximize the expected net benefit from approving and rejecting drugs. In our analysis, we consider the system in steady state and assume that the system parameters are unchanging in time.

Let B_{AE} denote the per-drug benefit associated with approving an effective drug, and let C_{AI} and C_{RE} denote the per-drug costs associated with approving ineffective (type I error) and rejecting effective (type II error) drugs, respectively. The per-drug benefit associated with rejecting an ineffective drug is normalized to zero. Each per drug benefit and cost is multiplied by the expected number of drugs $\mathbb{E}[N_{AE}(\alpha)]$, $\mathbb{E}[N_{AI}(\alpha)]$, and $\mathbb{E}[N_{RE}(\alpha)]$ in the post-review stages; expressions for these quantities for each model are given in Table 2. Let $\psi_i(\alpha) = \lambda_i(\alpha)/\mu_i$ where $i \in \{AE, AI, RE\}$ and $\lambda_i(\alpha)$ are given by equations (5)-(8).

Table 2 Expected number of drugs in post-review stages.

Model	$\mathbb{E}[N_{AE}(\alpha)]$	$\mathbb{E}[N_{AI}(\alpha)]$	$\mathbb{E}[N_{RE}(\alpha)]$
$M/M/\infty$	$\psi_{AE}(\alpha)$	$\psi_{AI}(\alpha)$	$\psi_{RE}(\alpha)$
$M/M/K/K$	$K\psi_{AE}(\alpha)\frac{\Gamma(K,\psi_{AE}(\alpha))}{\Gamma(K+1,\psi_{AE}(\alpha))}$	$\psi_{AI}(\alpha)$	$\psi_{RE}(\alpha)$

Note. $\Gamma(s,x) = (s-1)!e^{-x}\sum_{k=0}^{s-1}\frac{x^k}{k!}$ (where s is a positive integer) is the incomplete upper gamma function.

The optimal approval policy α^* is chosen to maximize the expected net benefit $V(\alpha)$:

$$\begin{aligned}\alpha^* &= \arg\max_{\alpha\in[0,1]} \; V(\alpha) \\ &= \arg\max_{\alpha\in[0,1]} \; \left\{B_{AE}\mathbb{E}[N_{AE}(\alpha)] - C_{AI}\mathbb{E}[N_{AI}(\alpha)] - C_{RE}\mathbb{E}[N_{RE}(\alpha)]\right\}.\end{aligned} \tag{9}$$

In the following analysis, $V_\infty(\alpha)$ and $V_K(\alpha)$ denote the expected net benefit for the $M/M/\infty$ and K-bumping models, respectively. Let α^*_∞ and V^*_∞ denote the optimal approval policy and the optimal expected net benefit under the $M/M/\infty$ model. Similarly, let α^*_K and V^*_K denote the optimal approval policy and the optimal expected net benefit under the K-bumping model.

4.2. Analysis of $M/M/\infty$ and K-Bumping Models

We first examine the structure of the optimal approval policies for each model, to gain insights into how the pre- and post-review characteristics of a drug impact the ultimate approval decision. All proofs are presented in Appendix A.

We first solve for the optimal approval policy under the $M/M/\infty$ model, which is unique.

THEOREM 1. *Under the $M/M/\infty$ model, the expected net benefit function $V_\infty(\alpha)$ is concave in the approval policy α. The optimal approval policy is given by*

$$\alpha^*_\infty = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}}\log\left(\frac{1-p}{p}\frac{C_{AI}/\mu_{AI}}{B_{AE}/\mu_{AE}+C_{RE}/\mu_{RE}}\right) + \frac{\delta\sqrt{I_n}}{2}\right). \tag{10}$$

The optimal approval policy α^*_∞ is increasing in the benefit B_{AE} of approving an effective drug and the cost C_{RE} of rejecting an effective drug, but is decreasing in the cost C_{AI} of approving an ineffective drug. This means that drugs with greater benefits (due to market size or disease specifics) or higher rejection costs (due to a type II error) should be easier to approve, while drugs with higher approval costs (due to a type I error) should be harder to approve. It is also easier to approve drugs with higher prior probabilities p of being effective, as they are more likely to offer a benefit to patients. Prolonging the durations of the approved effective $1/\mu_{AE}$ or rejected effective $1/\mu_{RE}$ stages increases the time that approved drugs can offer benefits, or rejected drugs incur costs, thus making drug approval more attractive. Increasing the duration of the approved ineffective stage $1/\mu_{AI}$ increases the time these ineffective drugs spend on the market, and thus should disincentivize approval.

REMARK 1. Under the $M/M/\infty$ model, the optimal approval policy α^*_∞ (10) does not depend on the R&D intensity λ, the rate of clinical trial completion μ_{CT}, and the rate of abandonment from clinical trials μ_{AB}.

To interpret this result, consider the NDA arrival rate $\tilde{\lambda} = \lambda\frac{\mu_{CT}}{\mu_{CT}+\mu_{AB}}$, which we denote as the NDA intensity. Note that it is sufficient to focus on this parameter because the pre-review parameters λ, μ_{CT}, and μ_{AB} appear in the expected net benefit function $V_\infty(\alpha)$ only through $\tilde{\lambda}$. The lack

of dependence of α^*_∞ on pre-review parameters stems from $\tilde{\lambda}$ having the same marginal effect on the costs and benefits of each post-review stage, when all stages are modeled as $M/M/\infty$ queues.

Under some specific assumptions, we obtain the approval policy derived by Montazerhodjat and Lo (2015). These authors study the FDA approval decision using a cost-minimization model, where they assign one-time costs to type I and type II errors, but do not consider the benefits of approving drugs. In contrast, our model incorporates both a per unit time error cost and the average duration over which this cost is incurred. By equating these two ways of measuring the cost of type I and type II errors, and by setting the benefit of approving an effective drug to zero, we obtain the same approval policy as Montazerhodjat and Lo (2015). We note that these authors frame the approval policy in terms of a threshold that the raw Wald statistic is compared to, rather than in terms of the p-value, but there is a one-to-one relationship between these two quantities. The following corollary summarizes this result.

COROLLARY 1. *If the total cost of a type I and type II error is defined as* $C_1 = C_{AI}/\mu_{AI}$ *and* $C_2 = C_{RE}/\mu_{RE}$*, respectively, and if there is no benefit from approving effective drugs* ($B_{AE} = 0$)*, then* α^*_∞ *reduces to*

$$\alpha^*_\infty = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}} \log\left(\frac{1-p}{p}\frac{C_1}{C_2}\right) + \frac{\delta\sqrt{I_n}}{2}\right).$$

We next analyze the more complex K-bumping model. In contrast to the simpler $M/M/\infty$ model, in which we analytically solve for the optimal policy, we find that there is no closed form solution for the optimal policy of the K-bumping model.

THEOREM 2. *For the K-bumping model, the expected net benefit function* $V_K(\alpha)$ *is concave in* α*. The optimal policy* α^*_K *satisfies the following fixed point equation:*

$$B_{AE}\frac{\partial \mathbb{E}[N_{AE}(\alpha^*_K)]}{\partial \psi_{AE}}\frac{\partial \psi_{AE}(\alpha^*_K)}{\partial \alpha} - C_{AI}\frac{\partial \psi_{AI}(\alpha^*_K)}{\partial \alpha} - C_{RE}\frac{\partial \psi_{RE}(\alpha^*_K)}{\partial \alpha} = 0. \tag{11}$$

From Theorem 2, we see that a unique approval policy exists that optimizes the expected net benefit function $V_K(\alpha)$ under the K-bumping model. Despite the lack of a closed form expression for the optimal policy, we can analyze the comparative statics of α^*_K with respect to the model parameters using the fixed point equation (11), as summarized in Proposition 1.

PROPOSITION 1. *The optimal approval policy* α^*_K *for the K-bumping model is*

(a) increasing in B_{AE}, C_{RE}, μ_{AI}, *and* μ_{AB},

(b) decreasing in C_{AI}, μ_{RE}, μ_{CT}, *and* λ,

(c) increasing in p *and decreasing in* μ_{AE} *under the additional assumption that*

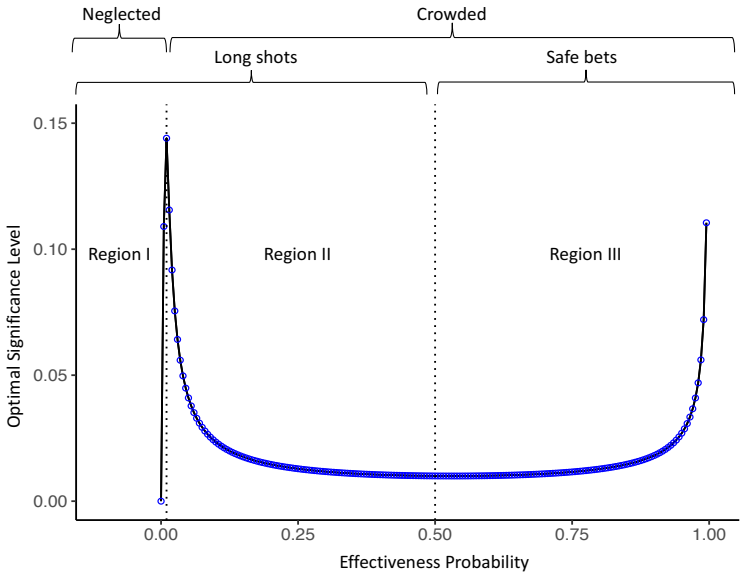
$$\psi_{AE}(\alpha^*_K) \leq -\frac{\frac{\partial \mathbb{E}[N_{AE}(\alpha^*_K)]}{\partial \psi_{AE}}}{\frac{\partial^2 \mathbb{E}[N_{AE}(\alpha^*_K)]}{\partial \psi^2_{AE}}}. \tag{12}$$

From Theorem 1 and Proposition 1, we see that the direction of the comparative statics for α_K^* with respect to B_{AE}, C_{RE}, C_{AI}, μ_{AI}, and μ_{RE} are the same as for α_∞^*, and thus the K-bumping model preserves the intuition about how these parameters impact the optimal approval policy. However, under the K-bumping model, the approval policy α_K^* now depends on the pre-review parameters. The optimal policy is decreasing in the the R&D intensity λ and the clinical trial rate μ_{CT}, so it is more difficult to approve drugs for diseases with many compounds in development and short clinical trial durations. The optimal policy is increasing in the abandonment rate μ_{AB}, so it is easier to approve drugs for diseases with high attrition rates.

Under the $M/M/\infty$ model, we show that as the prior probability p that drugs are effective increases, the optimal response is to approve more drugs. Similarly, we show that as the duration $1/\mu_{AE}$ of the market patent increases, the optimal policy approves more drugs. However, Proposition 1 implies that, under the K-Bumping model, there are situations in which the probability that drugs are effective p or time on the market $1/\mu_{AE}$ increases and yet the optimal response is to approve *fewer* drugs. To understand this counterintuitive result, consider the following example illustrated in Figure 3.

Suppose a particular disease has a high rate of R&D intensity $\tilde{\lambda}$, and high benefits of approving effective drugs B_{AE} relative to the costs of approving ineffective drugs C_{AI}. We divide Figure 3 into three regions to highlight how the drugs' characteristics non-monotonically drive the optimal

Figure 3 Example of the sensitivity of the optimal significance level (α_K^*) with respect to the effectiveness probability (p) if Proposition 1c is violated.



Note. $\sigma = 1$, $\delta = 0.10$, $n = 500$, $\tilde{\lambda} = 100$, $K = 10$, $B_{AE} = 1$, $\mu_{AE} = 0.10$, $C_{AI} = 0.01$, $\mu_{AI} = 0.10$, $C_{RE}/\mu_{RE} = 0$
Region I corresponds to $0 \leq p \leq 0.01$, Region II to $0.01 < p \leq 0.5$, and Region III to $0.5 < p \leq 1$.

policy. Each region can be understood in terms of two characteristics: the effectiveness probability p and the degree of crowding in the market for approved effective drugs, $\mathbb{E}[N_{AE}(\alpha)]$. In this example, we can think of drugs with a low probability ($p < 0.5$) of effectiveness as *long shots*, whereas those with high probability ($p \geq 0.5$) as *safe bets*. The market for approved effective drugs is *crowded* if many drugs are available (), or *neglected* if few are available. Region I is the intersection of a neglected market with long shot drugs. As the probability of being effective increases, the optimal policy is to approve more drugs (despite their low effectiveness probability) because of the large benefits of effective drugs and the paucity of drugs available for patients. In Region II, drugs are still long shots, but the market is crowded, so the optimal policy is to approve fewer drugs as the effectiveness probability increases. In a crowded market, approving additional drugs does not increase benefits—it only increases the rate of bumping. With long shot drugs, however, the high costs of approving ineffective drugs outweigh the expected benefits. Finally, in Region III, the market is crowded, but candidate drugs are reasonably safe bets, so the optimal policy is to approve more drugs as the effectiveness probability increases. Approving more drugs is optimal because safe bet drugs are likely effective and will bump other approved effective drugs, thereby avoiding type I error costs while marginally increasing benefits.

Condition (12) guarantees monotonicity of α_K^* with respect to p and $1/\mu_{AE}$ when $\psi_{AE}(\alpha_K^*) = \lambda_{AE}(\alpha_K^*)/\mu_{AE}$ is below a given bound. As $\psi_{AE}(\alpha_K^*)$ increases, the expected number of approved effective drugs $\mathbb{E}[N_{AE}(\alpha_K^*)]$ also increases, leading to increased rates of bumping. As demonstrated in Figure 3, high rates of bumping (due to a crowded market) contribute to the non-monotonicity of the optimal policy, and thus $\psi_{AE}(\alpha_K^*)$ must be bounded above in order for monotonicity to hold. The given upper bound can be interpreted using the concept of absolute risk aversion. An individual is said to be risk averse if, when faced with two investments with similar expected returns but different levels of risk, they will choose the investment with less risk. The absolute risk aversion of a utility function $u(x)$, defined as $-u''(x)/u'(x)$, measures how risk averse an individual is. The higher the curvature of $u(x)$, the higher the absolute risk aversion. Noting that the bound in (12) is the inverse of the absolute risk aversion of $\mathbb{E}[N_{AE}(\alpha_K^*)]$, we see that the higher the curvature of $\mathbb{E}[N_{AE}(\alpha_K^*)]$, the more sensitive the rate of bumping is to changes in $\psi_{AE}(\alpha_K^*)$ and thus the lower the bound on $\psi_{AE}(\alpha_K^*)$ must be to guarantee monotonicity. In other words, as drug markets become crowded more quickly—due to competing drugs gaining approval—the more likely the optimal policy is to exhibit counter-intuitive behavior with respect to the prior probability p and the market patent duration $1/\mu_{AE}$.

So far, we have analyzed the structure of the optimal approval policies for the $M/M/\infty$ and K-bumping models, as well as their sensitivity with respect to the model parameters. We conclude this section by comparing the optimal policies for these two models.

PROPOSITION 2. *The optimal approval policies satisfy*

$$\alpha_{LB} \leq \alpha_1^* \leq \cdots \leq \alpha_K^* \leq \cdots \leq \alpha_\infty^*$$

where

$$\alpha_{LB} = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}} \log\left(\frac{1-p}{p} \frac{C_{AI}/\mu_{AI}}{B_{AE}/(\mu_{AE}(1+\tilde{\lambda}p/\mu_{AE})^2) + C_{RE}/\mu_{RE}}\right) + \frac{\delta\sqrt{I_n}}{2}\right). \tag{13}$$

Furthermore, the optimal expected net benefit functions satisfy

$$V_1(\alpha_1^*) \leq \cdots \leq V_K(\alpha_K^*) \leq \cdots \leq V_\infty(\alpha_\infty^*).$$

Proposition 2 states that both the optimal approval policy and the optimal expected net benefit is non-decreasing in K. This result is intuitive, as increasing K increases the upper bound $B_{AE}K$ on the benefit associated with approved effective drugs, which makes approval more appealing. As K tends to infinity, we recover the $M/M/\infty$ model, and thus it is reasonable that α_∞^* and $V_\infty(\alpha_\infty^*)$ provide an upper bound for α_K^* and $V_K(\alpha_K^*)$, respectively. By analyzing the $M/M/1/1$ model, we are able to derive a non-trivial lower bound α_{LB} on any optimal approval policy. The bound closely resembles α_∞^*, with the only difference being that the rate μ_{AE} is replaced by $\mu_{AE}(1+\tilde{\lambda}p/\mu_{AE})^2$, which is a function of pre-approval parameters. This bound is particularly useful because the optimal approval policy for the K-bumping model has no closed form solution.

5. Numerical Study

To illustrate our queueing network model, we conduct a numerical study using publicly available drug approval data for three high-burden diseases: breast cancer, HIV, and hypertension. We compute the optimal approval policies for each disease under the $M/M/\infty$ and K-bumping models. The goals of this analysis are (i) to illustrate how the choice of model can result in different approval policies, (ii) to understand the impact of the development parameters on the optimal policy, and (iii) to demonstrate how our model can be used to develop disease-specific drug approval recommendations.

5.1. Parameter Estimation

We provide an overview of our model parameter estimation, with a detailed discussion and sources provided in Appendix B.

Clinical trial parameters. The values of n, λ, μ_{CT} and μ_{AB} are numerically estimated using clinical trial data from `clinicaltrials.gov`. We only include data from clinical trials that list the disease of interest as the primary condition in the trial, are categorized as interventional (rather than behavioral) studies, and use drug interventions (as opposed to medical devices). Clinical trial

enrollment n is estimated as the median enrollment of Phase III clinical trials, which are most often used to demonstrate efficacy. We used median enrollment, rather than the mean, due to a small number of trials with large enrollments. The rate λ at which drugs begin clinical trials is estimated using the mean number of trials that began for each year in the data. The rate μ_{CT} at which clinical trials are completed is estimated as the reciprocal of the sum of the mean durations of Phase I, Phase II, and Phase III trials. To estimate the rate of clinical trial abandonment μ_{AB}, note that $\mu_{AB} = \frac{\mu_{CT}[1-\mathbb{P}(\text{Complete clinical trials})]}{\mathbb{P}(\text{Complete clinical trials})}$. We first estimate the probability that a given drug intervention completes all three phases of clinical trials $\widehat{\mathbb{P}}(\text{Complete clinical trials})$. Given this estimate and our estimate $\widehat{\mu_{CT}}$ for the clinical trial completion rate, we obtain the following expression:

$$\widehat{\mu_{AB}} = \frac{\widehat{\mu_{CT}}\left[1-\widehat{\mathbb{P}}(\text{Complete clinical trials})\right]}{\widehat{\mathbb{P}}(\text{Complete clinical trials})}. \tag{14}$$

Due to a lack of publicly available data on individual drug effectiveness, we set the drug response standard deviation $\sigma = 1$, the treatment effect of a candidate drug $\delta = 0.125$, and the prior probability $p = 0.5$, the same values given in Montazerhodjat and Lo (2015). The choice of $p = 0.5$ assumes that drugs that have completed clinical trials and are facing FDA review are equally likely to be effective or ineffective. This assumption is conservative, likely underestimating the true efficacy of candidate drugs at this stage in the development process.

Number of drug classes. We identify classes of drugs that are widely recognized amongst health care providers. For example, classes for hypertension drugs include ACE Inhibitors, Beta Blockers, Calcium Channel Blockers, etc. Next, we use current treatment guidelines to remove classes rendered obsolete by newer therapies. For breast cancer, we use recommendations from the Quantum Leap Healthcare Collaborative and the National Comprehensive Cancer Network. For HIV, we use sources from the U.S. Department of Health and Human Services and the World Health Organization. For hypertension, we use sources from the Agency for Healthcare Research and Quality. Lists of all drug classes and references are provided in Appendix B.

Drug benefits and costs. We consider the benefit B_{AE} as the increase in overall U.S. life expectancy from approving an effective drug. We choose to measure benefits in terms of life expectancy because the three diseases considered are in the top 25 causes of premature mortality. Other measures, such as quality-adjusted life year (QALYs) could instead be used. To estimate this parameter, we examine the total change in U.S. life expectancy from the time that the first drug for each disease was approved to the current day, and attribute a portion of this change to advances in drug treatment of the disease. This increase in life expectancy is then divided by the number of drugs that have been approved to treat each disease to obtain a per drug estimate of the benefit associated with approving effective drugs.

In accordance with this interpretation of the per-drug benefit B_{AE}, we similarly interpret the costs C_{AI} and C_{RE} as the decreases in overall U.S. life expectancy from approving an ineffective or rejecting an effective drug, respectively. We view the total costs C_{AI}/μ_{AI} and C_{RE}/μ_{RE} as percentages ν_1 and ν_2, respectively, of the total benefit B_{AE}/μ_{AE}. To capture the notion that, by taking an ineffective drug, one is missing out on the treatment benefits of an effective drug, we conservatively use a value of $\nu_1 = 100\%$ in our numerical study. We view the total cost C_{RE}/μ_{RE} as the opportunity cost of rejecting an effective drug and use a value of $\nu_2 = 10\%$.

Post-review durations. We set the duration of the approved effective phase to reflect the 20 year patent life of approved drugs, which includes time on the market as well as time spent in clinical trials (patents are typically filed prior to the start of clinical trials). Thus we set μ_{AE} so that $1/\mu_{CT} + 1/\mu_{AE} = 20$ years.

5.2. Case Studies: Breast Cancer, HIV, and Hypertension

We conduct a numerical study of three high-burden diseases to illustrate the performance of our model. These diseases collectively account for over 10% of all drugs in development in 2016 (Murray 2013, PhRMA 2016a). Parameter estimates for each disease are summarized in Table 3, with additional details provided in Appendix B.

Breast cancer is an acute disease responsible for over 40,000 deaths in the U.S. each year (Breast Cancer Society 2016). Of the 250,000 new diagnoses each year, most patients complete primary treatment in the form of surgery, radiation, and chemotherapy within one year of diagnosis (Breast Cancer Society 2016). Additional hormone or targeted therapies may be prescribed for several years after primary treatment in order to reduce the risk of recurrence. Women living with metastatic breast cancer may take some form of oncological therapy for the remainder of their lives.

HIV, or Human Immunodeficiency Virus, is a virus that attacks the body's immune system, leaving individuals at risk for potentially deadly opportunistic infections. Patients typically take antiretroviral medications, which can suppress the amount of virus in the body, slow disease progression, and substantially prolong life (United States Department of Health and Human Services 2015). Currently 1.2 million Americans are living with HIV, and more than 12,000 die each year (Centers for Disease Control and Prevention 2016).

Hypertension, or high blood pressure, is a chronic condition, and diagnosed individuals take medications to control their blood pressure throughout their life. Hypertension affects 106 million Americans and is a precursor for heart disease, which is responsible for one in every four deaths (Hall et al. 2015, Centers for Disease Control and Prevention 2015).

Significant heterogeneity exists in the benefits (and costs, given our assumptions) associated with the three diseases. Hypertension has the highest per-drug total benefit B_{AE}/μ_{AE}, while HIV

Table 3 Parameter estimates for selected diseases.

Parameter	Breast Cancer	HIV	Hypertension	Source
n	647	354	418	clinicaltrials.gov
λ	10.94	5.41	4.94	clinicaltrials.gov
μ_{CT}	0.08	0.14	0.31	clinicaltrials.gov
μ_{AB}	0.46	0.28	0.20	clinicaltrials.gov
$\tilde{\lambda}$	1.64	1.80	2.99	clinicaltrials.gov
σ	1	1	1	Assumed
δ	0.125	0.125	0.125	Assumed
p	0.5	0.5	0.5	Assumed
$\delta\sqrt{I_n}$	2.25	1.66	1.81	clinicaltrials.gov
K	10	6	9	AHRQ, Drugs@FDA, DHHS, NCCN, QLHC, WHO
B_{AE}/μ_{AE}	0.019	0.009	0.596	B_{AE}: CDC.gov, Drugs@FDA, $1/\mu_{AE}+1/\mu_{CT}=20$
C_{AI}/μ_{AI}	0.019	0.009	0.596	Assumed
C_{RE}/μ_{RE}	0.002	0.001	0.060	Assumed

Sources: clinicaltrials.gov (National Library of Medicine and National Institutes of Health 2017); Agency for Healthcare Research and Quality (Townsend et al. 2011); Drugs@FDA (U.S. Food and Drug Administration 2017c); U.S. Department of Health and Human Services (2016); National Comprehensive Cancer Network (2016); Quantum Leap Healthcare Collaborative (2017); World Health Organization (2016); Centers for Disease Control and Prevention (2017)

medications have the lowest per-drug benefits. This effect is driven by the difference in disease prevalence and the shorter duration of clinical trials for hypertension, permitting these medications to spend more time on the market before patent expiry. Breast cancer and hypertension both have large numbers of drug classes, while HIV has relatively few. The small number of HIV drug classes may, in part, result from the first case of HIV in the United States being identified in 1981, while treatments for hypertension and breast cancer have been in development since the 1950s (U.S. Department of Health and Human Services 2016).

These diseases also differ with respect to the pre-review process. Breast cancer has the highest R&D intensity λ, but also the highest clinical trial attrition rate μ_{AB}, resulting in an NDA intensity $\tilde{\lambda}$ of 1.64 drugs per year. According to Arrowsmith and Miller (2013), this high rate of attrition stem from difficulty in establishing efficacy for oncology drugs in trials with relatively short durations. In contrast, hypertension has the lowest R&D intensity λ at 4.94 drugs per year, but also the lowest attrition rate, leading to the highest NDA intensity of 2.99 drugs per year.

Using the parameter values given in Table 3, we calculate the optimal approval policies for the three diseases under both the $M/M/\infty$ model and the K-bumping model. Our results, summarized in Table 4, highlight the differences between these diseases as well as the contrast between our two models. The optimal approval policy of the K-bumping model for hypertension is below the 5% traditional significance level, while all other approval policies exceed this value.

In concordance with Proposition 2, the policies found using the $M/M/\infty$ model approve more drugs than those of the K-bumping model. The K-bumping model, whose optimal policy depends on the pre-review process, recommends making breast cancer drugs the easiest to approve due to

Table 4 **Optimal approval policies and percentage change in expected approval benefits, expected approval costs, and expected rejection costs compared to the traditional $\alpha = 5\%$ threshold for 3 high burden diseases.**

	Optimal Policy		%Δ Expected Approval Benefits		%Δ Expected Approval Costs		%Δ Expected Rejection Costs	
	α^*_∞	α^*_K	$M/M/\infty$	K-Bumping	$M/M/\infty$	K-Bumping	$M/M/\infty$	K-Bumping
Breast Cancer	0.140	0.126	+19%	+17%	+179%	+152%	-55%	-51%
HIV	0.220	0.072	+16%	+5%	+339%	+44%	-62%	-14%
Hypertension	0.197	0.042	+5%	-1%	+295%	-17%	-61%	+8%

their low NDA intensity $\tilde{\lambda}$ and hypertension drugs the hardest to approve due to their high NDA intensity. In contrast, the $M/M/\infty$ model policy recommends making HIV drugs the easiest to approve and breast cancer drugs the hardest due to differences in clinical trial information $\delta\sqrt{I_n}$. As we use a value of $\delta = 0.125$ for all diseases, differences in clinical trial information are driven by differences in trial enrollments n. Under our cost assumptions, high clinical trial information leads to more stringent approval policies because high information improves the ability to statistically differentiate between ineffective and effective drugs. Imposing stricter approval policies reduces type I error costs while maintaining high benefits.

Table 4 also shows how the expected approval benefits $B_{AE}\mathbb{E}[N_{AE}(\alpha)]$, expected approval costs $C_{AI}\mathbb{E}[N_{AI}(\alpha)]$, and expected rejection costs $C_{RE}\mathbb{E}[N_{RE}(\alpha)]$ under both queueing models compare to a traditional significance level of $\alpha = 5\%$. Note that all policies, with the exception of the optimal policy of the K-bumping model for hypertension, exceed the $\alpha = 5\%$ significance level. The $M/M/\infty$ model recommends less stringent approval policies than the K-bumping model, generating greater increases in expected approval benefits, but also higher approval costs due to type I errors. The difference between the $M/M/\infty$ and K-bumping models is particularly pronounced in the case of HIV. Compared to the $\alpha = 5\%$ policy, the $M/M/\infty$ policy for HIV therapies increases approval benefits by 16% and approval costs by 339%, while the K-bumping HIV policy results in a smaller gain (5%) in benefits but also a substantially smaller (44%) increase in approval costs. On the rejection side, $M/M/\infty$ policies reject fewer drugs than with K-bumping, and thus incur fewer rejection costs.

We conduct sensitivity analysis of the optimal approval policies for both queueing models with respect to the nominal parameter values, displayed in Table 3. The purpose of this analysis is (i) to identify which parameters have the strongest effect on the optimal approval policy and (ii) to understand the impact of parameter uncertainty on the optimal policy.

We focus on four key parameters for our analysis: the NDA intensity $\tilde{\lambda}$, the benefit B_{AE} of approving effective drugs, the duration $1/\mu_{AE}$ of the approved effective stage, and the prior probability p. The parameter $\tilde{\lambda}$ was chosen because it summarizes the pre-review parameters λ, μ_{CT}, and μ_{AB} and because although the $M/M/\infty$ policy does not depend on $\tilde{\lambda}$, the K-bumping model

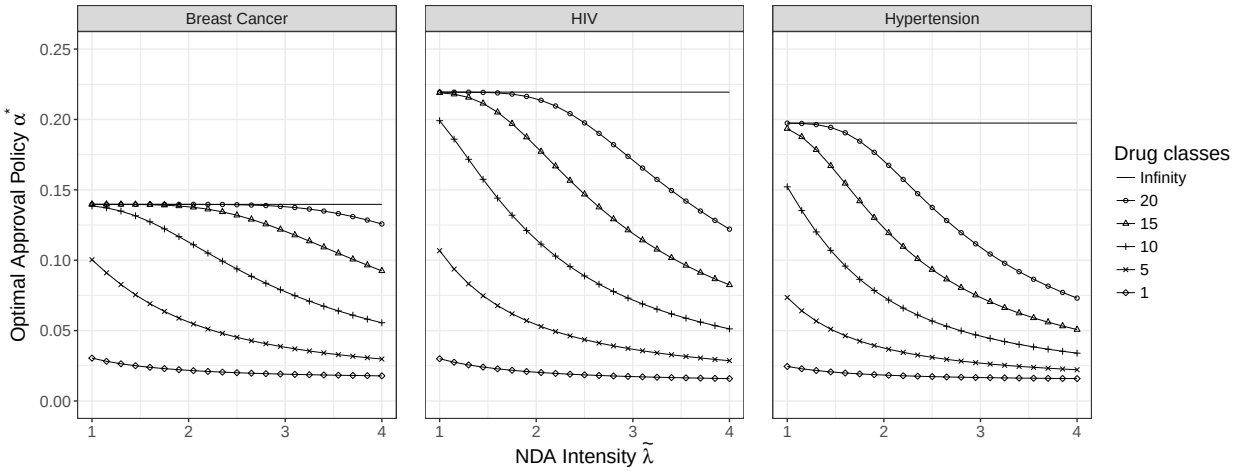
does. The benefit B_{AE} was chosen to illustrate the importance of this parameter in determining the optimal policy. The approved effective stage duration $1/\mu_{AE}$ and the prior probability p were chosen in order to further explore the non-monotonic nature of the relationship between these parameters and the optimal policy under the K-bumping model.

We vary $\tilde{\lambda}$ from 1 drug per year to 4 drugs per year, B_{AE} from no increase in overall U.S. life expectancy to a 0.1 year increase, $1/\mu_{AE}$ from 0 years to 20 years, and p from 0 to 1. The range for p is natural, the range for $1/\mu_{AE}$ covers the possible durations of the market patent, and the ranges for the other parameters are chosen to demonstrate the range of values that the optimal policy can take.

For each parameter, and for each disease, we plot the value of the optimal policies for both the $M/M/\infty$ and K-bumping models, where the number of drug classes $K = 1, 5, 10, 15$, and 20. In the resulting plots, shown in Figures 4-7, we see that the approval policies computed from the K-bumping model are increasing in K and bounded above by the approval policy of the $M/M/\infty$ model; this observation is supported by Proposition 2.

Sensitivity with respect to NDA intensity. As suggested by Theorem 1 and Proposition 1, Figure 4 shows that the optimal approval policy α^*_∞ does not depend on the NDA intensity $\tilde{\lambda}$, while the optimal policy for the K-bumping model α^*_K is decreasing in this parameter. We also observe that the α^*_K curves switch from concave to convex as $\tilde{\lambda}$ increases, and as K increases so does the value of $\tilde{\lambda}$ at which this transition occurs. The change in concavity can be explained by the rapid increase in the rate of bumping of approved effective drugs that occurs as more drug candidates enter the NDA process. As more bumping occurs, approving drugs becomes increasingly risky because drugs are either effective, in which case they do not generate much marginal benefit, or they are ineffective and generate additional costs. The optimal policy given increased rates of bumping is to approve fewer drugs to avoid these potential costs. For small values of K, there is

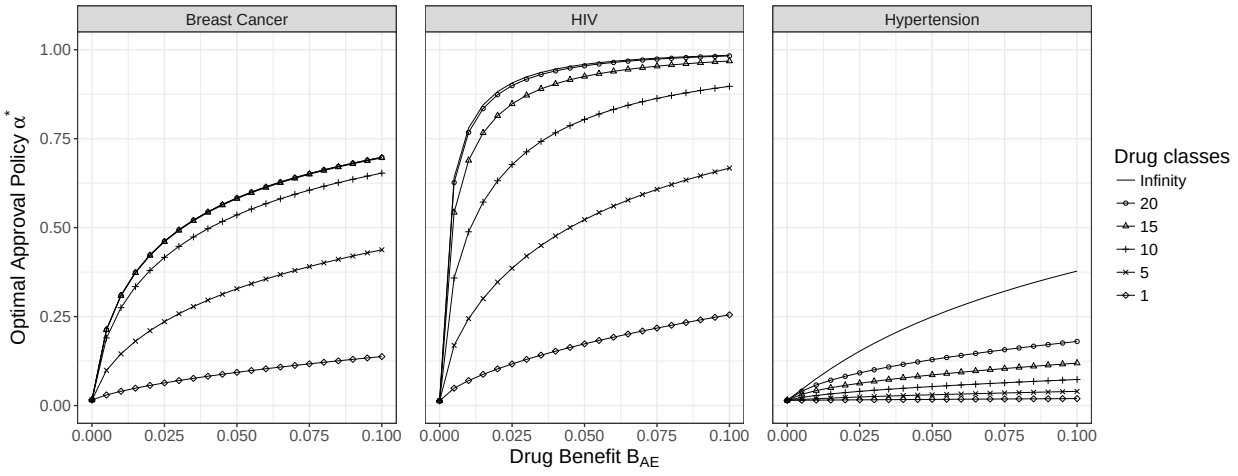
Figure 4 Sensitivity of the optimal approval policy to the NDA intensity $\tilde{\lambda}$



less space on the market for approved effective drugs, so bumping occurs even for low values of $\tilde{\lambda}$, causing the α_K^* curves to decrease at an increasing rate (switch to convex). In contrast, for large values of K, there is more market opportunity and bumping only occurs for higher values of $\tilde{\lambda}$, meaning that the concavity switch occurs for larger values of $\tilde{\lambda}$.

Sensitivity with respect to drug benefit. The optimal policy curves for both models are concave increasing in the benefit B_{AE} of approving effective drugs, indicating that the optimal policy has decreasing marginal returns with respect to B_{AE}, as shown in Figure 5. Our estimates for B_{AE} range from 0.0007 for HIV to 0.0355 for hypertension, as B_{AE} is the per drug benefit per year on the market. One potential concern is that the approval policies presented in Table 4 could change substantially if our estimates for B_{AE} are biased. Based on the results of Lichtenberg (2002), who estimates an average gain in life expectancy due to drug therapies across all diseases is 0.016 per drug per year, we believe that, if anything, our estimates are biased slightly downwards and our resulting policies are more conservative. We also note that the optimal policy becomes increasingly sensitive to B_{AE} as the number of drug classes K increases, and is the most sensitive for the $M/M/\infty$ model. When K is small, an increase in B_{AE} has a smaller impact on the total benefit $B_{AE}\mathbb{E}[N_{AE}(\alpha)]$—and thus a smaller impact on the optimal policy— than when K is large.

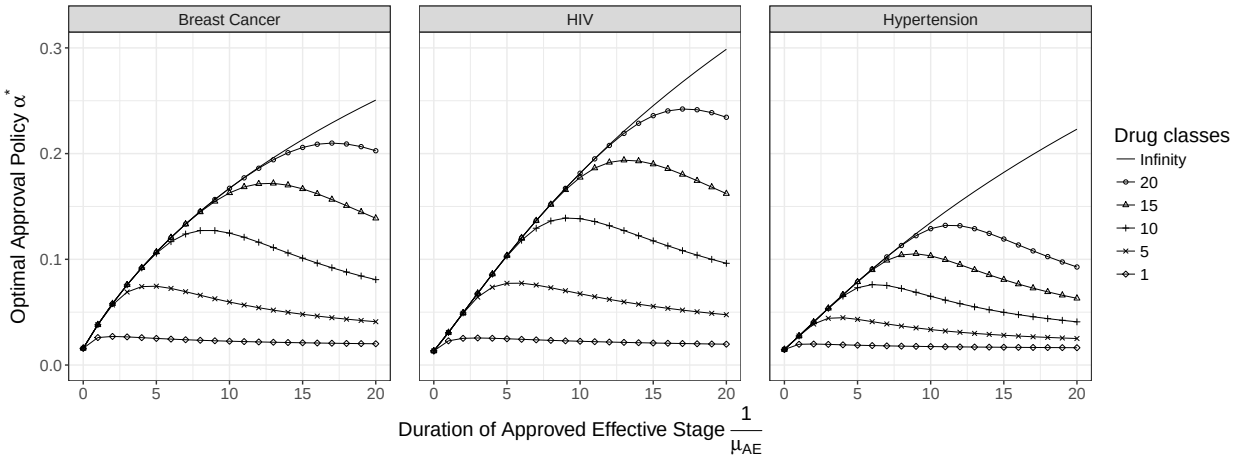
Figure 5 Sensitivity of the optimal approval policy to the drug benefit B_{AE}



Sensitivity with respect to market patent duration. Recall from Theorem 1 and Proposition 1 that the optimal policy α_∞^* is monotonically increasing in the market patent duration $1/\mu_{AE}$, but that the optimal policy α_K^* is non-monotonic in this parameter. This behavior, more prominently seen for larger values of K in Figure 6, in which the α_K^* curves first increase and then decrease, can be explained by the bumping of approved effective drugs. Recall that high rates of bumping result in more conservative approval policies because approved drugs do not generate additional benefit if effective and may result in costs if ineffective. When the market patent duration $1/\mu_{AE}$ is short, the expected number of approved effective drugs $\mathbb{E}[N_{AE}(\alpha)]$ is small and the

rate of bumping is low. As $1/\mu_{AE}$ increases, the benefits of approving additional drugs outweigh the costs because there is still room on the market for approved effective drugs, so the optimal policy α_K^* increases. However, as $1/\mu_{AE}$ continues to increase, more approved effective drugs are on the market and the rate of bumping increases to the point where the benefits of approving drugs no longer outweigh the costs, resulting in a decrease in the optimal policy.

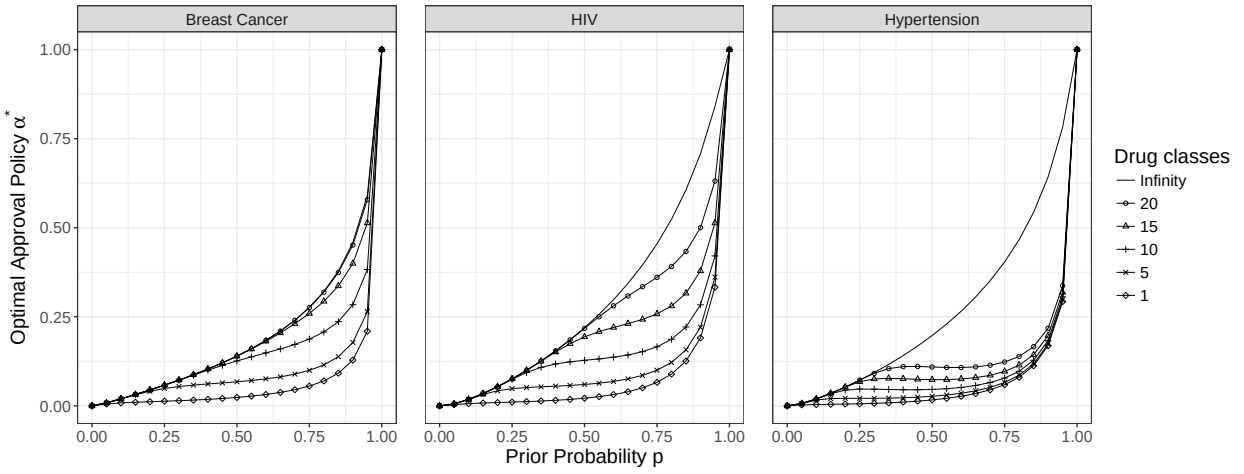
Figure 6 Sensitivity of the optimal approval policy to the market patent duration $1/\mu_{AE}$



Sensitivity with respect to prior probability. As shown in Figure 7, the optimal approval policy for both models approves no drugs when $p = 0$, and approves all drugs when $p = 1$. For the $M/M/\infty$ model, increasing p always increases the optimal approval policy, but this is not always true for the K-bumping model. The plot for hypertension most clearly illustrates the non-monotonic relationship between p and α_K^*; the possibility of observing this behavior was demonstrated via the example in Figure 3 in which condition (12) in Proposition 1 was violated. The approval policy α_K^* is sensitive to changes in p when p is close to 1, but is less sensitive to lower values of p. As we assume a uniform prior $p = 0.5$ in our case study, this observation suggests that our estimates of the optimal policies for the three diseases are robust to small changes in p.

In addition to the sensitivity analysis in the preceding sections, we also analyze the factors ν_1 and ν_2, where $C_{AI}/\mu_{AI} = \nu_1 B_{AE}/\mu_{AE}$ and $C_{RE}/\mu_{RE} = \nu_2 B_{AE}/\mu_{AE}$, varying each parameter from 0 to 2. Note that when computing the optimal policies, we assume values $\nu_1 = 1$ and $\nu_2 = 0.10$. We find that the optimal policy is very sensitive to the value of ν_1 when $\nu_1 < 1$ and becomes less sensitive as ν_1 increases, while the optimal policy is comparatively less sensitive to ν_2. Changing the type I error costs has a large impact on the optimal policy because these are the only costs that act to discourage drug approval, while changing the type II error costs has less of an impact because both benefits and type II error costs act to encourage drug approval.

Figure 7 Sensitivity of the optimal approval policy to the prior probability of effectiveness p



6. Discussion

Our proposed queueing framework offers several insights into the FDA drug approval process, demonstrating how the pre-review process and the existing market saturation could influence the disease-specific optimal FDA approval policy. Our model captures three barriers to having many effective drugs available to treat diseases: (i) low innovation in new drug formulation (i.e., a low *arrival rate*), (ii) lengthy clinical trials (i.e., a low *service rate*), and (iii) high rates of attrition in the development process (i.e., a high *abandonment rate*). Over the years, the FDA has introduced a variety of programs designed to address these challenges. Our model could be used to evaluate the effectiveness of these programs, in terms of their impact on societal health benefits. For example, the *Breakthrough Therapy* program, which expedites the development and review of drugs that demonstrate substantial improvement over available therapy, could be analyzed by studying the effect of decreasing the duration of the clinical trials stage in our queueing model.

Not only could our framework be used to evaluate the impact of bringing drugs to market more quickly, but it offers a fundamentally different way of analyzing the drug approval problem. For example, the FDA's *orphan drug designation* policy aims to mitigate the shortage of research funding allotted to rare diseases by providing incentives, such as tax credits for clinical trial testing, to companies that develop treatments for these conditions. Another way of addressing low research intensity, which is explored in our work, is to incentivize drug developers by easing approval standards for diseases with few drugs in development (i.e., a low arrival rate). Our framework could provide solutions to more challenging problems within the development process. For example, one contributing factor for the high attrition rate observed for certain diseases is a lack of understanding of the underlying disease mechanism. Adjusting FDA approval standards based on disease-specific attrition rates could help address this issue.

In addition to considering characteristics of the development process, our work recommends adjusting approval standards based on aspects of the existing market. More specifically, we consider the role of drug obsolescence, in which old drugs are replaced by newer compounds, and substitution occurs amongst available treatments. If there is a high degree of obsolescence or drug substitution for a particular disease (i.e., a low value of K), then approval standards should be stricter, as the market cannot support a large number of drugs. We measure obsolescence as the number of unique drug classes available to treat each disease, but a variety of other measures could be used. For example, one could consider the number of drugs that exceed a given market share (e.g. 10%) for a disease, or the number of drugs typically administered during a course of treatment.

Our model relates to the Bayesian Decision Analysis approach used by Montazerhodjat and Lo (2015), including our use of a similar drug efficacy framework and post-review cost structure. Unlike this study, however, we estimate a benefit of approving effective drugs. Using a queueing network model, we also explicitly integrate the pre-review process as well as obsolescence amongst drugs within the same class, two key aspects of the drug approval process that Montazerhodjat and Lo (2015) do not capture. As a result, the optimal approval policies differ substantially between studies. For example, they determine an approval policy of 1.35% for breast cancer, whereas we find an optimal approval policy of 12.6% (under our K-bumping model), primarily due to breast cancer's low rate of NDA submission. In contrast, our optimal policy for hypertension is 4.2%, a more conservative threshold than the 8% policy suggested by Montazerhodjat and Lo (2015).

6.1. Limitations

Our study has several limitations. First, drug efficacy is based on a single quantitative endpoint resulting from a balanced, two-arm randomized clinical trial. Modern trial designs are often unbalanced, have more than two arms, and involve multiple endpoints. Our model could be easily adapted for unbalanced trials, but incorporating multiple arms and endpoints would require a more sophisticated statistical framework and queueing model. With breast cancer, for example, potential clinical endpoints include the tumor size and time until recurrence, and it is unclear how these endpoints should be collectively used to establish drug efficacy. However, such disease-specific complexity could render our model analytically intractable.

Second, we make several simplifying assumptions regarding the FDA's decision making process. We do not consider qualitative aspects, such as concerns over clinical trial design nor labeling or manufacturing capabilities, as possible reasons for denying approval. We also do not consider that the FDA may request that a firm revise and resubmit an NDA, which occurs in about 30% of reviews (Downing et al. 2014). Additionally, we assume that the NDA filing and FDA review stage occur immediately; in reality, these reviews take six months, on average. Our model could be extended to incorporate such complexities, which would not likely change our main results.

We make several assumptions when computing the expected net benefit. First, we assume that all queues are in steady state and the number of drug classes K is fixed, rather than using a transient analysis and allowing K to vary with time. The assumption that a queueing system is in steady state is commonly used because transient analysis is often intractable. Next, we measure the per-drug benefit B_{AE} as the increase in overall U.S. life expectancy associated with approving an effective drug. Because breast cancer, HIV, and hypertension are amongst the top 25 causes of premature mortality in the U.S., measuring benefits in terms of life expectancy is reasonable. However, for diseases that have a substantial impact on quality of life, QALYs might provide a more appropriate measure of the benefits from drug treatment. Finally, we assume a linear relationship between the number of approved effective drugs and the expected total benefit, whereas diminishing marginal returns to approving additional effective drugs might exist. When estimating the benefit B_{AE} of approving an effective drug, however, we find that the cumulative gain in life expectancy associated with treating the three diseases fit a linear trend well, suggesting that a linear (or perhaps piecewise linear) relationship may be a good approximation (see Figure B1).

6.2. Future Work

Our study could motivate several directions for future work. One extension would be to analyze the drug development process using a game theoretic approach, with the FDA and a pharmaceutical company as players. Conditions under which a pharmaceutical company should conduct additional clinical trials and resubmit a rejected NDA, or when they should abandon the failed drug and begin developing a new product, could be explored. Our queueing model framework could also be used to study product development, in general. A firm developing multiple products that require similar resources as inputs faces the decision of which product(s) to invest in; a queueing model could capture the resulting delays in development. Upon entering the market, products may compete for market share, which could be modeled using a variant of our K-bumping model.

6.3. Conclusions

Faced with regulating thousands of drugs in a nation where millions are afflicted with severe diseases and advances in medical treatment have improved the quality and length of life, the FDA must find the correct balance between ensuring the safety and efficacy of drugs while spurring development of novel drugs and bringing life-saving products to market in a timely fashion. Our work offers a transparent, quantitative framework that can be used to assess candidate drugs based on severity and prevalence as well as characteristics of the drug development process and existing market. Such a model could augment the complex decision-making and statistical analyses conducted by the FDA, providing a more customized approach to policy-making.

References

Adler PS, Mandelbaum A, Nguyen V, Schwerer E (1995) From Project to Process Management: An Empirically-based Framework for Analyzing Product Development Time. *Management Science* 41(3):458–484.

Ahuja V, Birge JR (2016) Response-Adaptive Designs for Clinical Trials: Simultaneous Learning from Multiple Patients. *European Journal of Operational Research* 248(2):619–633.

AidsInfo (2017) FDA-Approved HIV Medicines. URL `https://aidsinfo.nih.gov/education-materials/fact-sheets/21/58/fda-approved-hiv-medicines`, Accessed Jan 2017.

Arrowsmith J, Miller P (2013) Trial Watch: Phase II and Phase III Attrition Rates 2011 to 2012. *Nature Reviews Drug Discovery* 12(8):569–569.

Bertsimas D, Hair AO, Relyea S, Silberholz J (2016) An Analytics Approach to Designing Combination Chemotherapy Regimens for Cancer. *Working Paper* .

Bertsimas D, Johnson M, Kallus N (2015) The Power of Optimization Over Randomization in Designing Experiments Involving Small Samples. *Operations Research* 63(4):868–876.

Breast Cancer Society (2016) Breast Cancer Fact Sheet. Technical report, URL `http://www.breastcancer.org/about_us/press_room/press_kit/facts_figures`, Accessed Mar 2017.

Casella G, Berger RL (2002) *Statistical Inference* (Thomson Learning), Second edition.

CDC (2014) United States Life Tables. Technical report.

CDER, CBER (1998) Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. Technical report.

Centers for Disease Control and Prevention (2015) Heart Disease Facts. URL `https://www.cdc.gov/heartdisease/facts.htm`, Accessed Mar 2017.

Centers for Disease Control and Prevention (2016) HIV in the United States: At a Glance. URL `https://www.cdc.gov/hiv/statistics/overview/ataglance.html`, Accessed Mar 2017.

Centers for Disease Control and Prevention (2017) Compressed Mortality. URL `https://wonder.cdc.gov/mortSQL.html`, Accessed Feb 2017.

Code of Federal Regulations (2016) Title 21, Volume 5, Subpart E. URL `https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=312&showFR=1&subpartNode=21:5.0.1.1.3.5`, Accessed Jun 2017.

Cossin D, Schellhorn H (2007) Credit Risk in a Network Economy. *Management Science* 53:1604–1617.

DiMasi JA, Hansen RW, Grabowski HG (2003) The Price of Innovation: New Estimates of Drug Development Costs. *Journal of Health Economics* 22(2):151–185.

Ding M, Eliashberg J (2002) Structuring the New Product Development Pipeline. *Management Science* 48(3):343–363.

Dohrman AJ (2005) Rethinking and Restructuring the FDA Drug Approval Process in Light of the Vioxx Recall. *The Journal of Corporation Law* .

Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS (2014) Clinical Trial Evidence Supporting FDA Approval of Novel Therapeutic Agents, 2005-2012. *JAMA* 311(4):368.

DrugWatch (2017) Vioxx Lawsuit. URL `https://www.drugwatch.com/vioxx/lawsuit/`, Accessed Mar 2017.

Friedman LM, Furberg CD, DeMets DL (2015) *Fundamentals of Clinical Trials, 4th ed.* (Springer Science and Business Media), 5th edition.

Google, World Bank (2017) United States Life Expectancy. URL `https://goo.gl/ffvyzW`, Accessed Jan 2017.

Govil MK, Fu MC (1999) Queueing Theory in Manufacturing: A Survey. *Journal of Manufacturing Systems* 18(3):214–240.

Green L (2006) Queueing Analysis in Healthcare. Hall RW, ed., *Patient Flow: Reducing Delay in Healthcare Delivery*, 281–307 (Boston, MA: Springer US).

Hall HI, An Q, Tang T, Song R, Chen M, Green T, Kang J (2015) Prevalence of Diagnosed and Undiagnosed HIV Infection–United States, 2008-2012. Technical report.

Harel A (1990a) Convexity Properties of the Erlang Loss Formula. *Operations Research* 38(3):499–505.

Harel A (1990b) Convexity Properties of the Erlang Loss Formula. *Operations Research* 38(3):499–505.

Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J (2014) Clinical Development Success Rates for Investigational Drugs. *Nature Biotechnology* 32(1):40–51.

IMS Health (2016) Medicines Use and Spending in the U.S. A Review of 2015 and Outlook to 2020. Technical report.

Jennison C, Turnbull W (2000) Two-Sided Tests : Introduction. *Group Sequential Methods with Applications to Clinical Trials* (Chapman and Hall/CRC).

Kaplan EH (2010) Terror Queues. *Operations Research* 58(4):773–784.

Killen CP, Hunt Ra, Kleinschmidt EJ (2007) Managing the New Product Development Project Portfolio: A Review of the Literature and Empirical Evidence. *Portland International Center for Management of Engineering and Technology* 1864–1874.

Kinch MS, Haynesworth A, Kinch SL, Hoyer D (2014) An Overview of FDA-Approved New Molecular Entities: 1827-2013. *Drug Discovery Today* 19(8):1033–1039.

Krishnan V, Ulrich KT (2001) Product Development Decisions: A Review of the Literature. *Management Science* 47(1):1–21.

Lichtenberg FR (2002) Sources of U.S. Longevity Increase, 1960-1997. *Working Paper* .

Montazerhodjat V, Lo A (2015) Is the FDA Too Conservative or Too Aggressive?: A Bayesian Decision Analysis of Clinical Trial Design. *Working Paper* .

Murray CJL (2013) The State of US Health, 1990-2010. *JAMA* 310(6):591.

National Cancer Institute (2016) Drugs Approved for Breast Cancer. URL `https://www.cancer.gov/about-cancer/treatment/drugs/breast`, Accessed Jan 2017.

National Comprehensive Cancer Network (2016) NCCN Guidelines for Patients, Stage 0-Stage IV. Technical report.

National Library of Medicine, National Institutes of Health (2017) CinicalTrials.gov. URL `https://clinicaltrials.gov/ct2/home`, Accessed Jan 2017.

Öner KB, Kiesmüller GP, van Houtum GJ (2009) Monotonicity and Supermodularity Results for the Erlang Loss System. *Operations Research Letters* 37(4):265–268.

PhRMA (2013) 2013 Biopharmaceutical Research Industry Profile. Technical report.

PhRMA (2015a) Biopharmaceutical Research and Development: The Process Behind New Medicines. Technical report.

PhRMA (2015b) Medicines in Development for Cancer. Technical report.

PhRMA (2016a) 2016 Profile: Biopharmaceutical Research Industry. Technical report.

PhRMA (2016b) Medicines in Development for Alzheimer's Disease. Technical report.

Quantum Leap Healthcare Collaborative (2017) Breast Cancer Drugs. URL `https://www.breastcancertrials.org/BCTIncludes/Resources/BreastCancerDrugs.html`, Accessed Jan 2017.

Sterne JaC, Smith GD (2001) Sifting the Evidence - What's Wrong with Significance Tests? *Physical Therapy* 322(8):226–231.

Townsend R, Leonard C, Lopez de Nava K (2011) Utilization of Antihypertensive Drug Classes Among Medicare Beneficiaries with Hypertension, 2007-2009. Technical report, Agency for Healthcare Research and Quality.

Tufts Centre for the Study of Drug Development (2014) Cost of Developing a New Drug Innovation in the Pharmaceutical Industry. Technical report.

United States Department of Health and Human Services (2015) Health, United States, 2015 with Special Feature on Racial and Ethnic Health Disparities. Technical report.

US Department of Health and Human Services (2016) A Timeline of HIV/AIDS. URL `https://www.aids.gov/hiv-aids-basics/hiv-aids-101/aids-timeline/`, Accessed Jan 2017.

US Food and Drug Administration (2014) The FDA's Drug Review Process: Ensuring Drugs are Safe and Effective. URL `http://www.fda.gov/Drugs/ResourcesForYou/Consumers/ucm143534.htm`, Accessed Mar 2017.

US Food and Drug Administration (2015) Fast Track, Breakthrough Therapy, Accelerated Approval, Priority Review. URL http://www.fda.gov/forpatients/approvals/fast/ucm20041766.htm, Accessed Mar 2017.

US Food and Drug Administration (2016a) Antiretroviral Drugs Used in the Treatment of HIV Infection. URL https://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm, Accessed Jan 2017.

US Food and Drug Administration (2016b) Drug Safety and Availability. Lotronex (Alosetron Hydrochloride) Information.

US Food and Drug Administration (2016c) FDA Adverse Event Reporting System (FAERS). URL http://www.fda.gov/Drugs/InformationOnDrugs/ucm135151.htm, Accessed Mar 2017.

US Food and Drug Administration (2016d) Frequently Asked Questions on Patents and Exclusivity. URL http://www.fda.gov/Drugs/DevelopmentApprovalProcess/ucm079031.htm, Accessed Mar 2017.

US Food and Drug Administration (2017a) Designating an Orphan Product: Drugs and Biological Products. URL http://www.fda.gov/ForIndustry/DevelopingProductsforRareDiseasesConditions/HowtoapplyforOrphanProductDesignation/default.htm, Accessed Mar 2017.

US Food and Drug Administration (2017b) Drug Recalls. URL http://www.fda.gov/Drugs/Drugsafety/DrugRecalls/default.htm, Accessed Feb 2017.

US Food and Drug Administration (2017c) Drugs@FDA. URL https://www.accessdata.fda.gov/scripts/cder/daf/, Accessed Jan 2017.

US Food and Drug Administration (2017d) Enhancing Benefit-Risk Assessment in Regulatory Decision-Making. URL https://www.fda.gov/forindustry/userfees/prescriptiondruguserfee/ucm326192.htm, Accessed Aug 2017.

World Health Organization (2016) Consolidated Guidelines on HIV Prevention, Diagnosis, Treatment and Care for Key Populations. Technical report.

Yapar O, Gans N, Chick SE (2016) Proposal for Fully Sequential Multiarm Trials with Correlated Arms. *Proceedings of the 2016 Winter Simulation Conference*, 3688–3689.

Appendix A: Proofs

We introduce the following notation and equations to be used throughout this section. We suppress the dependence of various terms on α for readability and only explicitly note it when needed for clarity. At times, we may use the notation $\mathbb{E}[N_{AE}^K]$ in regards to the K-bumping model to stress the dependence of this term on K. For all derivatives, the variable of differentiation is α unless otherwise specified.

Recall that $\psi_i(\alpha) = \lambda_i(\alpha)/\mu_i$ for $i \in \{AE, AI, RE\}$, where $\lambda_i(\alpha)$ are given in equations (5)-(8).

The expected net benefit and its derivative with respect to α for the $M/M/\infty$ model are given by

$$\begin{aligned} V_\infty(\alpha) &= B_{AE}\psi_{AE}(\alpha) - C_{AI}\psi_{AI}(\alpha) - C_{RE}\psi_{RE}(\alpha) \qquad \text{(A.15)} \\ &= \tilde{\lambda}\left(\frac{B_{AE}}{\mu_{AE}} p\left(1-\Phi\left(\Phi^{-1}(1-\alpha)-\delta\sqrt{I_n}\right)\right) - \frac{C_{AI}}{\mu_{AI}}(1-p)\alpha - \frac{C_{RE}}{\mu_{RE}} p\,\Phi\left(\Phi^{-1}(1-\alpha)-\delta\sqrt{I_n}\right)\right) \end{aligned}$$

$$V_\infty'(\alpha) = B_{AE}\frac{\partial\psi_{AE}(\alpha)}{\partial\alpha} - C_{AI}\frac{\partial\psi_{AI}(\alpha)}{\partial\alpha} - C_{RE}\frac{\partial\psi_{RE}(\alpha)}{\partial\alpha} \qquad \text{(A.16)}$$

The expected net benefit and its derivative with respect to α for the K-bumping model are given by

$$V_K(\alpha) = B_{AE}K\psi_{AE}(\alpha)\frac{\Gamma(K,\psi_{AE}(\alpha))}{\Gamma(K+1,\psi_{AE}(\alpha))} - C_{AI}\psi_{AI}(\alpha) - C_{RE}\psi_{RE}(\alpha) \qquad \text{(A.17)}$$

$$V_K'(\alpha) = B_{AE}\frac{\partial\mathbb{E}[N_{AE}^K(\alpha)]}{\partial\psi_{AE}}\frac{\partial\psi_{AE}(\alpha)}{\partial\alpha} - C_{AI}\frac{\partial\psi_{AI}(\alpha)}{\partial\alpha} - C_{RE}\frac{\partial\psi_{RE}(\alpha)}{\partial\alpha} \qquad \text{(A.18)}$$

Note that the expected net benefit functions (A.15) and (A.17) are only defined for $\alpha \in [0,1]$.

To simplify the exposition of some proofs, we introduce the Erlang B function, defined as

$$B(K,\psi_{AE}) = \mathbb{P}(N_{AE} = K) = \frac{\psi_{AE}^K/K!}{\sum_{j=0}^K \psi_{AE}^j/j!} \qquad \text{(A.19)}$$

The expected number of approved effective drugs in the K-bumping model can be defined either in terms of the incomplete upper gamma function Γ or in terms of the Erlang B function as follows:

$$\begin{aligned} \mathbb{E}[N_{AE}^K(\alpha)] &= K\psi_{AE}\frac{\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})} \qquad \text{(A.20)} \\ &= (1-B(K,\psi_{AE}))\psi_{AE} \qquad \text{(A.21)} \end{aligned}$$

Proof of Theorem 1:

To show that $V_\infty(\alpha)$ is concave, we compute the second derivative of (A.15) with respect to α and observe that this quantity is clearly negative for all values of α:

$$V_\infty''(\alpha) = -\tilde{\lambda}\left(\frac{B_{AE}}{\mu_{AE}} + \frac{C_{RE}}{\mu_{RE}}\right) e^{\Phi^{-1}(1-\alpha)\delta\sqrt{I_n} - \delta^2 I_n/2 + (\Phi^{-1}(1-\alpha))^2/2} < 0$$

Thus, setting (A.16) equal to zero and solving for α results in expression (10) for α_∞^*. Note that $\alpha_\infty^* \in [0,1]$ because $\Phi \in [0,1]$. ■

Proof of Theorem 2: Under the K-bumping model, the number of drugs in the approved effective post-approval stage behaves according to a Markov chain. Let p_j be the stationary probability that there are j drugs in this stage. A straightforward application of the flow balance equations gives

$$p_j = \frac{\psi_{AE}^j/j!}{\sum_{i=0}^K \psi_{AE}^i/i!} \qquad \forall j = 0,\ldots,K$$

These are the same stationary probabilities as the $M/M/K/K$, or Erlang loss, system. We use this result to obtain the expression (A.17) for the expected net benefit function.

To show that $V_K(\alpha)$ is concave in α, we argue that $B_{AE}K\psi_{AE}\frac{\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})}$, $-C_{AI}\psi_{AI}$, and $-C_{RE}\psi_{RE}$ are all concave functions of α, and thus the sum of concave functions is concave. It is straightforward to establish that ψ_{AE} is concave in α, and thus by Lemma 1, $B_{AE}K\psi_{AE}\frac{\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})}$ can be expressed as the composition of a concave increasing function with a concave function and is thus concave. Establishing concavity of $C_{AI}\psi_{AI}$ and $-C_{RE}\psi_{RE}$ is similarly straightforward. We note that in the case that $\alpha > 0$, $-C_{RE}\psi_{RE}$ and $C_{AI}\psi_{AI}$ are strictly concave in α and thus so is $V_K(\alpha)$.

The dependence of α_K^* on the pre-review parameters μ_{CT}, μ_{AB}, and λ is established in the proof of Proposition 1. ■

LEMMA 1. *$\mathbb{E}[N_{AE}^K]$ is a non-decreasing function of ψ_{AE}.*

Proof of Lemma 1: Consider the first derivative of $\mathbb{E}[N_{AE}^K]$ with respect to ψ_{AE}:

$$\frac{\partial \mathbb{E}[N_{AE}^K]}{\partial \psi_{AE}} = \frac{e^{-\psi_{AE}}\psi_{AE}^{K+1}\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})^2} + \frac{\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})} - \frac{e^{-\psi_{AE}}\psi_{AE}^K}{\Gamma(K+1,\psi_{AE})} \tag{A.22}$$

Harel (1990a) proved that the expected number of busy servers in an Erlang Loss system is concave in the offered load rate, which establishes that (A.22) is decreasing in ψ_{AE}. We demonstrate that $\lim_{\psi_{AE}\to\infty}\frac{\partial \mathbb{E}[N_{AE}^K]}{\partial \psi_{AE}} = 0$. This, combined with the fact that (A.22) is decreasing in ψ_{AE} implies that $\frac{\partial \mathbb{E}[N_{AE}^K]}{\partial \psi_{AE}}$ is non-negative, which gives the desired result.

$$\begin{aligned}
&\lim_{\psi_{AE}\to\infty} \frac{\partial \mathbb{E}[N_{AE}^K]}{\partial \psi_{AE}} \\
&= \lim_{\psi_{AE}\to\infty} \frac{e^{-\psi_{AE}}\psi_{AE}^K}{\Gamma(K+1,\psi_{AE})}\left(\psi_{AE}\frac{\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})} - 1\right) + \frac{\Gamma(K,\psi_{AE})}{\Gamma(K+1,\psi_{AE})} \\
&= \lim_{\psi_{AE}\to\infty} \frac{\psi_{AE}^K}{K!\sum_{j=0}^K \psi_{AE}^j/j!}\left(\frac{\sum_{j=0}^{K-1}\psi_{AE}^{j+1}/j!}{K\sum_{j=0}^K \psi_{AE}^j/j!} - 1\right) + \frac{\sum_{j=0}^{K-1}\psi_{AE}^j/j!}{K\sum_{j=0}^K \psi_{AE}^j/j!} \\
&= \lim_{\psi_{AE}\to\infty} \frac{\psi_{AE}^K}{K!\sum_{j=0}^K \psi_{AE}^j/j!} \cdot \lim_{\psi_{AE}\to\infty}\left(\frac{\sum_{j=0}^{K-1}\psi_{AE}^{j+1}/j!}{K\sum_{j=0}^K \psi_{AE}^j/j!} - 1\right) + \lim_{\psi_{AE}\to\infty}\frac{\sum_{j=0}^{K-1}\psi_{AE}^j/j!}{K\sum_{j=0}^K \psi_{AE}^j/j!} \\
&= 1\cdot(1-1) + 0 \quad ■
\end{aligned}$$

Proof of Proposition 1: By the Implicit Function Theorem, we have that

$$\frac{\partial \alpha_K^*}{\partial x} = -\frac{\frac{\partial V_K'(\alpha_K^*)}{\partial x}}{\frac{\partial V_K'(\alpha_K^*)}{\partial \alpha}}$$

where x is the parameter of interest. The fact that $V_K(\alpha)$ is concave in α means the denominator is negative and thus the sign of $\frac{\partial \alpha^*}{\partial x}$ is given by the sign of $\frac{\partial V_K'(\alpha_K^*)}{\partial x}$. We use equation (A.18) to find the sign of the effect of each parameter on α_K^*:

- $\operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial B_{AE}}\right) = \operatorname{sgn}\left(\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}}\frac{\partial \psi_{AE}(\alpha_K^*)}{\partial \alpha}\right) > 0$
- $\operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial C_{AI}}\right) = \operatorname{sgn}\left(-\frac{\partial \psi_{AI}(\alpha_K^*)}{\partial \alpha}\right) < 0$
- $\operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial C_{RE}}\right) = \operatorname{sgn}\left(-\frac{\partial \psi_{RE}(\alpha_K^*)}{\partial \alpha}\right) > 0$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial \mu_{AI}}\right) = \operatorname{sgn}\left(-C_{AI}\frac{\partial^2 \psi_{AI}(\alpha_K^*)}{\partial \alpha \partial \mu_{AI}}\right) > 0$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial \mu_{RE}}\right) = \operatorname{sgn}\left(-C_{RE}\frac{\partial^2 \psi_{RE}(\alpha_K^*)}{\partial \alpha \partial \mu_{RE}}\right) < 0$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial \tilde{\lambda}}\right) = \operatorname{sgn}\left(B_{AE}\left(\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}}\frac{\partial^2 \psi_{AE}(\alpha_K^*)}{\partial \alpha \partial \tilde{\lambda}} + \frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2}\frac{\partial \psi_{AE}(\alpha_K^*)}{\partial \tilde{\lambda}}\frac{\partial \psi_{AE}(\alpha_K^*)}{\partial \alpha}\right)\right.$$
$$\left. -C_{AI}\frac{\partial^2 \psi_{AI}(\alpha_K^*)}{\partial \alpha \partial \tilde{\lambda}} - C_{RE}\frac{\partial^2 \psi_{RE}(\alpha_K^*)}{\partial \alpha \partial \tilde{\lambda}}\right)$$

Multiplying both sides by $\tilde{\lambda} > 0$ (which does not change the sign) gives

$$\begin{aligned}\operatorname{sgn}\left(\tilde{\lambda}\frac{\partial \alpha_K^*}{\partial \tilde{\lambda}}\right) &= \operatorname{sgn}\left(B_{AE}\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}}\frac{\partial \psi_{AE}(\alpha_K^*)}{\partial \alpha} - C_{AI}\frac{\partial \psi_{AI}(\alpha_K^*)}{\partial \alpha} - C_{RE}\frac{\partial \psi_{RE}(\alpha_K^*)}{\partial \alpha}\right.\\ &\quad \left. + B_{AE}\frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2}\psi_{AE}(\alpha_K^*)\frac{\partial \psi_{AE}(\alpha_K^*)}{\partial \alpha}\right)\\ &= \operatorname{sgn}\left(B_{AE}\frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2}\psi_{AE}(\alpha_K^*)\frac{\partial \psi_{AE}(\alpha_K^*)}{\partial \alpha}\right) < 0\end{aligned}$$

The second equality is due to the first order condition for α_K^*. The sign of the last expression is negative due to the concavity of $\mathbb{E}[N_{AE}^K]$ with respect to ψ_{AE} and the fact that ψ_{AE} is increasing in α.

We claim that $\frac{\partial \alpha_K^*}{\partial \mu_{AE}}$ and $\frac{\partial \alpha_K^*}{\partial p}$ are non-monotonic and that (12) is a sufficient condition to ensure that $\frac{\partial \alpha_K^*}{\partial \mu_{AE}} \leq 0$ and $\frac{\partial \alpha_K^*}{\partial p} \geq 0$. The proof of this is given by straightforward calculation of the sign of the desired derivatives:

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial \mu_{AE}}\right) = \operatorname{sgn}\left(-\frac{\tilde{\lambda}}{\mu_{AE}^2}pe^{\Phi^{-1}(1-\alpha_K^*)\delta\sqrt{I_n}-\frac{\delta^2 I_n}{2}}\left(\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}}\right.\right.$$
$$\left.\left. + \frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2}\frac{\tilde{\lambda}}{\mu_{AE}}p\left[1-\phi\left(\phi^{-1}(1-\alpha_K^*)-\delta\sqrt{I_n}\right)\right]\right)\right)$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha_K^*}{\partial p}\right) = \operatorname{sgn}\left(\tilde{\lambda}e^{\Phi^{-1}(1-\alpha_K^*)\delta\sqrt{I_n}-\frac{\delta^2 I_n}{2}}\left[\frac{B_{AE}}{\mu_{AE}}\left(\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}}\right.\right.\right.$$
$$\left.\left.\left. + \frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2}\frac{\tilde{\lambda}}{\mu_{AE}}p\left[1-\Phi\left(\Phi^{-1}(1-\alpha_K^*)-\delta\sqrt{I_n}\right)\right]\right) + \frac{C_{RE}}{\mu_{RE}}\right] + \tilde{\lambda}\frac{C_{AI}}{\mu_{AI}}\right)$$

Note that these derivatives may be negative due to the terms $\frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2} < 0$. However, under (12) we have that

$$\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}} + \frac{\partial^2 \mathbb{E}[N_{AE}^K(\alpha_K^*)]}{\partial \psi_{AE}^2}\frac{\tilde{\lambda}}{\mu_{AE}}p\left(1-\Phi\left(\Phi^{-1}(1-\alpha_K^*)-\delta\sqrt{I_n}\right)\right) \geq 0$$

which implies $\frac{\partial \alpha^*}{\partial \mu_{AE}} \leq 0$ and $\frac{\partial \alpha^*}{\partial p} \geq 0$. Using condition (12) and recognizing that $\psi_{AE} = \frac{\tilde{\lambda}}{\mu_{AE}}p\left(1-\Phi\left(\Phi^{-1}(1-\alpha_K^*)-\delta\sqrt{I_n}\right)\right)$, we see that the above inequality is positive. ■

Proof of Proposition 2: We begin by demonstrating that $\alpha_1^* \leq \alpha_2^* \leq \cdots \leq \alpha_K^*$. To do this, we show that $V_K'(\alpha_{K+1}^*) \leq 0$ for any $K \geq 1$. The concavity of $V_K(\alpha)$ will imply the desired inequality. Consider the following expression:

$$V_K'(\alpha_{K+1}^*) - V_{K+1}'(\alpha_{K+1}^*) = B_{AE}\frac{\partial \psi_{AE}(\alpha_{K+1}^*)}{\partial \alpha}\left(\frac{\partial \mathbb{E}[N_{AE}^K(\alpha_{K+1}^*)]}{\partial \psi_{AE}} - \frac{\partial \mathbb{E}[N_{AE}^{K+1}(\alpha_{K+1}^*)]}{\partial \psi_{AE}}\right) \tag{A.23}$$

From the optimality of α_{K+1}^*, we know that $V_{K+1}'(\alpha_{K+1}^*) = 0$. By Lemma 2, $\frac{\partial \mathbb{E}[N_{AE}^K]}{\partial \psi_{AE}}$ is non decreasing in K, and thus we see that $V_K'(\alpha_{K+1}^*) \leq 0$. As this holds for any K, we obtain the desired result. To obtain a lower

bound α_{LB} on all optimal policies, we find a lower bound on α_1^* and use the fact that $\alpha_1^* \leq \alpha_K^*$ for all K. We find that α_1^* must satisfy the following equation:

$$\alpha_1^* = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}}\log\left(\frac{1-p}{p}\frac{C_{AI}/\mu_{AI}}{B_{AE}/(\mu_{AE}(1+\psi_{AE}(\alpha_1^*))^2)+C_{RE}/\mu_{RE}}\right)+\frac{\delta\sqrt{I_n}}{2}\right) \tag{A.24}$$

Let us denote the right hand side of (A.24) by $f(\alpha)$. We see that $f(\alpha)$ is decreasing in α and that α_1^* satisfies $\alpha_1^* = f(\alpha_1^*)$. Thus $\alpha_1^* \geq f(\alpha)$ for all $\alpha \geq \alpha_1^*$. In particular, we see that $\alpha_1^* \geq f(1) = \alpha_{LB}$, which establishes the desired lower bound (13).

We now demonstrate that $\alpha_K^* \leq \alpha_\infty^*$. Once again, we use the concavity of $V_K(\alpha)$ to establish the result. Consider the following expression:

$$V_K'(\alpha_\infty^*) - V_\infty'(\alpha_\infty^*) = B_{AE}\frac{\partial\psi_{AE}(\alpha_\infty^*)}{\partial\alpha}\left(\frac{\partial\mathbb{E}[N_{AE}^K(\alpha_\infty^*)]}{\partial\psi_{AE}(\alpha_\infty^*)} - 1\right)$$

By the optimality of α_∞^*, we have that $V_\infty'(\alpha_\infty^*) = 0$, and thus $V_K'(\alpha_\infty^*) \leq 0$ iff $\frac{\partial\mathbb{E}[N_{AE}^K]}{\partial\psi_{AE}(\alpha)}$ is less than or equal to 1. To see this is indeed the case, note that $\frac{\partial\mathbb{E}[N_{AE}^K]}{\partial\psi_{AE}(\alpha)}$ is decreasing in $\psi_{AE}(\alpha)$ – by Harel (1990b) – and is equal to 1 when $\psi_{AE}(\alpha) = 0$. The fact that this term is positive and bounded above by 1 gives that $V_K'(\alpha_\infty^*) \leq 0$, establishing that $\alpha_K^* \leq \alpha_\infty^*$.

Next, we show that $V_K(\alpha_K^*) \leq V_{K+1}(\alpha_{K+1}^*)$. To do this, we first show that $V_K(\alpha) \leq V_{K+1}(\alpha)$ for all α. To see that this holds, note that the only term in the expected net benefit function that depends on K is $\mathbb{E}[N_{AE}^K]$, and recall that this term can be written in terms of the Erlang B function as (A.21). Using the fact that the Erlang B function is non-increasing in K, it is clear that the expected number of approved effective drugs is non-decreasing in K. From this, we have that $V_K(\alpha) \leq V_{K+1}(\alpha)$. The series of inequalities $V_K(\alpha_K^*) \leq V_{K+1}(\alpha_K^*) \leq V_{K+1}(\alpha_{K+1}^*)$ completes the proof.

Finally, we demonstrate that $V_K(\alpha_K^*) \leq V_\infty(\alpha_\infty^*)$. To do this, we first show that $V_K(\alpha) \leq V_\infty(\alpha)$ for all α. This is equivalent to showing

$$V_\infty(\alpha) - V_K(\alpha) = B_{AE}\psi_{AE}(\alpha)\left(1 - K\frac{\Gamma(K,\psi_{AE}(\alpha))}{\Gamma(K+1,\psi_{AE}(\alpha))}\right) \geq 0$$

The desired result follows from the inequality $1 \geq K\frac{\Gamma(K,\psi_{AE}(\alpha))}{\Gamma(K+1,\psi_{AE}(\alpha))}$, whose proof is straightforward.

The remainder of the proof follows from the series of inequalities $V_K(\alpha_K^*) \leq V_\infty(\alpha_K^*) \leq V_\infty(\alpha_\infty^*)$. ■

LEMMA 2. $\frac{\partial\mathbb{E}[N_{AE}^K]}{\partial\psi_{AE}}$ *is non-decreasing in* K.

Proof of Lemma 2: Recall that $\mathbb{E}[N_{AE}^K]$ can be defined in terms of the Erlang B function as (A.21). With this definition, we have

$$\frac{\partial\mathbb{E}^K[N_{AE}]}{\partial\psi_{AE}} = 1 - \psi_{AE}\frac{\partial B(K,\psi_{AE})}{\partial\psi_{AE}} - B(K,\psi_{AE})$$

Using the fact that

$$\frac{\partial B(K,\psi_{AE})}{\partial\psi_{AE}} = B(K,\psi_{AE})\left(\frac{K}{\psi_{AE}} - 1 + B(K,\psi_{AE})\right)$$

the condition that $\frac{\partial\mathbb{E}[N_{AE}^K]}{\partial\psi_{AE}}$ is increasing in K can be written as follows:

$$\frac{\partial}{\partial\psi_{AE}}\left[\psi_{AE}\left(B(K,\psi_{AE}) - B(K+1,\psi_{AE})\right)\right] \geq 0$$

This is a statement that the function $\psi_{AE}\left(B(K,\psi_{AE}) - B(K+1,\psi_{AE})\right)$ is increasing in ψ_{AE}. This function has been studied extensively in queueing theory and is known as *the load carried by the last server*. The fact that this function is increasing in ψ_{AE} for all K was established in Öner et al. (2009). ■

Appendix B: Parameter Estimation

Clinical trial parameters. For each of the diseases (breast cancer, HIV, and hypertension), we perform an Advanced Search on clinicaltrials.gov with the following field settings:

`Search Terms: (insert disease here)`

`Study Type: Interventional Studies`

`Conditions: (insert disease here)`

`Interventions: Drug`

All other field settings were left blank. After downloading the data that resulted from this search, we remove trials that met the following exclusion criterion: (i) Non-drug intervention (Behavioral, Biological, Device, Dietary Supplement, Other, Procedure, Genetic, Radiation), (ii) Conditions other than the disease of interest, (iii) Enrollment $= 0$ or NULL, (iv) Study Completion Date or Study Start Date NULL, (v) Duration of study $= 0$ or NULL, (vi) Study Start Date before January 2000 or Study Completion Date after January 2017, (vii) Title or Condition fields do not indicate relevance of the trial to the disease of interest, (viii) Drug listed in intervention was not related to treating the disease of interest. Using the trial data that remain after imposing exclusion criterion (i)-(viii), we estimate the following parameters.

- **Clinical trial enrollment**. We use the median enrollment for Phase III trials, as these are the trials that are most often used for demonstrating efficacy of the candidate drug. We choose to use median enrollment rather than mean enrollment as the distribution of enrollment is skewed to the right by a small number of trials with extremely large enrollments.

- **Rate of new drug development.** We group all trials associated with a given drug intervention. For each intervention, we define a variable `Year of First Trial` to be the year in which the first trial (either Phase I, Phase II, or Phase III) for that intervention began. We allow the first trial to be a trial of any phase, rather than requiring the first trial to be a Phase I trial because pharmaceutical companies are not required to register Phase I trials with clinicaltrials.gov. Thus calculating the rate of new drug development as the number of Phase I trials initiated each year would underestimate the true rate. After calculating the year of first trial for each intervention, we calculate λ as the mean number of drug interventions that begin each year during the period 2000-2016.

- **Rate of clinical trial completion.** Rather than calculating the rate at which clinical trials are completed, we calculate $1/\mu_{CT}$, or the mean duration of clinical trials. Let D_i denote the mean duration of Phase i trials, where $i = 1, 2, 3$. We estimate $1/\mu_{CT}$ as $D_1 + D_2 + D_3$.

- **Rate of abandonment.** Recall for a drug that $\mathbb{P}(\text{complete clinical trials}) = \frac{\mu_{CT}}{\mu_{CT}+\mu_{AB}}$. For each drug intervention in our data, we define a binary variable `Completed Phase III` to be one if there is a Phase III or Phase IV trial associated with that intervention, and zero otherwise. Let $\widehat{\mathbb{P}}(\text{Complete clinical trials})$, our estimate of the probability of completing clinical trials, be the average of `Completed Phase III` across all drug interventions. Our estimate for μ_{AB} is given by:

$$\widehat{\mu_{AB}} = \frac{\widehat{\mu_{CT}}\left[1 - \widehat{\mathbb{P}}(\text{Complete clinical trials})\right]}{\widehat{\mathbb{P}}(\text{Complete clinical trials})}.$$

Number of drug classes. We use multiple medical, pharmaceutical, and FDA resources to construct a list of common classes of drugs and to determine whether or not each class of drugs is considered to be the standard of care. Drug classes that are not the standard of care are removed from our list. Table B2 gives the classifications (and thus the unique number of drug classes K) for breast cancer, HIV, and hypertension.

Drug benefit. Our procedure for calculating B_{AE} is to first estimate the portion of the increase in U.S. life expectancy that can be associated to reductions in mortality from treating the disease, and then divide this increase in life expectancy by the number of drugs that have been approved to treat the disease.

We use life expectancy estimates from 1900-2003 from the CDC National Vital Statistics Report (CDC 2014). Estimates for the years 2004-2015 were obtained from the World Bank (Google and World Bank 2017). We obtain mortality rate data from 1968-2015 (rate per 100,000 individuals) for the entire population as well as for each disease from the Wonder Compressed Mortality website (Centers for Disease Control and Prevention 2017). The ICD codes used for each diseases are given in Table B1.

For each year between 1968-2015, we calculate the percent of the total mortality rate that corresponds to each disease. Next, starting with 1969, we calculate the change in U.S. life expectancy compared to the previous year. We then multiply the percent of total mortality corresponding to each disease by the change in life expectancy. This is the proportion of the change in life expectancy for each year that we attribute to a given disease. Finally, we calculate the cumulative change in life expectancy attributed to each disease from 1969-2015. This gives us a proxy for the impact of treatment of the disease on life expectancy over time. We then create a plot of this cumulative gain from 1969-2015 and use linear regression to fit a trend line. Figure B1 shows the cumulative life expectancy gains and linear trend lines (shown as dashed lines) for all diseases. The R^2 for the breast cancer, HIV, and hypertension regressions are 0.9761, 0.8942, and 0.9312, respectively. Using this regression, we calculate the total gain in life expectancy by subtracting the cumulative gain in life expectancy in 2015 from the cumulative gain in the year that the first drug to treat the disease was approved by the FDA. Note that in some cases, the first drug was approved by the FDA prior to 1968. Finally, we divide this quantity by the number of FDA approved drugs for the disease to obtain our estimate of B_{AE}. The list of FDA approved drugs for breast cancer, HIV, and hypertension used in our calculations are given in Tables B3, B4, and B5, respectively.

Figure B1 Cumulative gain in life expectancy from disease treatment.

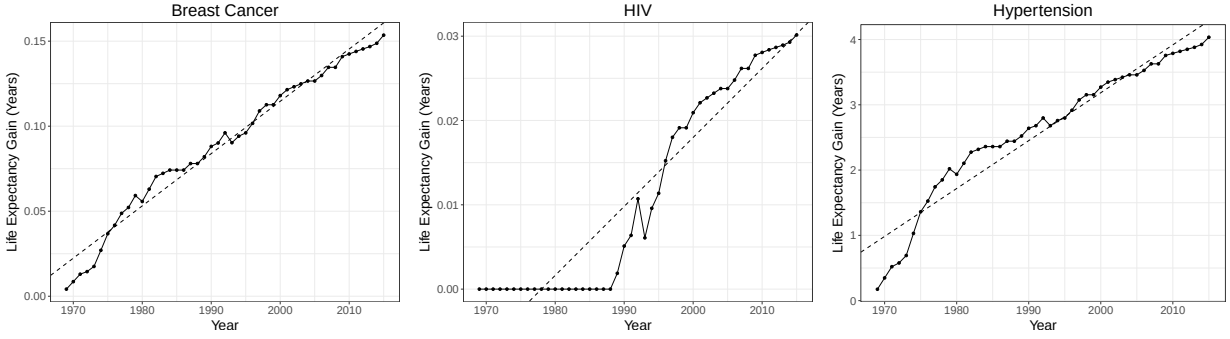


Table B6 lists the drugs that were initially approved, and later withdrawn, for the treatment of breast cancer, HIV, and hypertension. We provide this list for completion and to illustrate the small number of withdrawn drugs for each disease.

Table B1 ICD codes used to calculate mortality rates.

Disease	Time Period	Code (Description)
Breast cancer	1968–1978 (*ICD-8 Codes*)	174 (Malignant neoplasm of breast)
	1979–1988 (*ICD-9 Codes*)	174.0 (Nipple and areola)
		174.1 (Central portion)
		174.2 (Upper-inner quadrant)
		174.3 (Lower-inner quadrant)
		174.4 (Upper-outer quadrant)
		174.5 (Lower-outer quadrant)
		174.6 (Axillary tail)
		174.8 (Other specified sites of female breast)
		174.9 (Breast (female), unspecified)
		175 (Malignant neoplasm of male breast)
	1989–2015 (*ICD-10 Codes*)	C50.0 (Nipple & areola, malignant neoplasms)
		C50.1 (Central portion of breast, malignant neoplasms)
		C50.2 (Upper-inner quadrant of breast, malignant neoplasms)
		C50.3 (Lower-inner quadrant of breast, malignant neoplasms)
		C50.4 (Upper-outer quadrant of breast, malignant neoplasms)
		C50.5 (Lower-outer quadrant of breast, malignant neoplasms)
		C50.6 (Axillary tail of breast, malignant neoplasms
		C50.8 (Overlapping lesion of breast, malignant neoplasms)
		C50.9 (Breast, unspecified, malignant neoplasms)
HIV	1979–1988 (*ICD-9 Codes*)	042-044 (HIV infection)
	1989–2015 (*ICD-9 Codes*)	B20-B24 (HIV disease)
Hypertension	1968–1978 (*ICD-8 Codes*)	390-458 (Diseases of the circulatory system)
	1979–1988 (*ICD-9 Codes*)	390-459 (Diseases of the circulatory system)
	1989–2015 (*ICD-10 Codes*)	I00-I99 (Diseases of the circulatory system)

Source: Centers for Disease Control and Prevention (2017)

Table B2 Drug classifications by disease.

Disease	Drug Class	Source
Breast cancer	Alkylating Agents	QLHC (2017), NCCN (2016)
	Anthracyclines	QLHC (2017), NCCN (2016)
	Anti-Estrogen Drugs	QLHC (2017), NCCN (2016)
	Aromatase Inhibitors	QLHC (2017), NCCN (2016)
	Combination Chemo	QLHC (2017), NCCN (2016)
	Ovarian Suppression	QLHC (2017), NCCN (2016)
	Platinum Drugs	QLHC (2017)
	Targeted Biological Therapy (HER-2)	QLHC (2017), NCCN (2016)
	Taxanes	QLHC (2017)
	Vinca Agents	QLHC (2017)
HIV	Combination Therapy	DHHS (2016)
	Integrase Inhibitors	WHO (2016)
	Non-Nucleoside Reverse Transcriptase Inhibitors	WHO (2016)
	Nucleoside Reverse Transcriptase Inhibitors	WHO (2016)
	Pharmacokinetic Enhancers	DHHS (2016)
	Protease Inhibitors	WHO (2016)
Hypertension	Angiotensin Converting Enzyme (ACE) Inhibitors	AHRQ (2011)
	Angiotensin II Receptor Blockers (ARB)	AHRQ (2011)
	Antiadrenergics	AHRQ (2011)
	Beta Blockers	AHRQ (2011)
	Calcium Channel Blockers	AHRQ (2011)
	Combination Products	AHRQ (2011)
	Diuretics	AHRQ (2011)
	Other Renin-Angiotensin System Antagonists	AHRQ (2011)
	Vasodilators	AHRQ (2011)

Sources: Quantum Leap Healthcare Collaborative (2017); National Comprehensive Cancer Network (2016); U.S. Department of Health and Human Services (2016); World Health Organization (2016); Agency for Healthcare Research and Quality (Townsend et al. 2011).

Table B3 FDA-approved breast cancer drugs.

Drug (Brand Name)	Approval	Drug Class
Thiotepa (Tepadina)	March 1959	Alkylating Agents
Cyclophosphamide (Cytoxan)	May 2008	
Methotrexate (Trexall)	Aug 1959	Other Chemotherapy
Vinblastine (Velban)	Aug 1987	
Vincristine (Oncovin)	Apr 1988	
Fluorouracil 5-FU (Adrucil)	Aug 1991	
Gemcitabine (Gemzar)	May 1996	
Irinotecan (Camptosar)	Jun 1996	
Capecitabine (Xeloda)	Apr 1998	
Temozolomide (Temodar)	Aug 1999	
Ixabepilone (Ixempra)	Oct 2007	
Eribulin (Halaven)	Nov 2010	
Topotecan (Hycamtin)	Dec 2010	
Megestrol Acetate (Megace)	Aug 1971	Other Hormone Therapy
Cisplatin (Platinol)	Dec 1978	Platinum Drugs
Carboplatin (Paraplatin)	Mar 1989	
Goserelin (Zoladex)	Dec 1989	Ovarian Suppression
Leuprolide (Lupron)	Apr 1993	
Abarelix (Plenaxis)	Nov 2003	
Buserelin (Suprefact)	N/A	
Paclitaxel (Taxol)	Dec 1992	Taxanes
Docetaxel (Taxotere)	May 1996	
Paclitaxel (Abraxane)	Jan 2005	
Vinorelbine (Navelbine)	Dec 1994	Vinca Agents
Toremifine (Fareston)	May 1997	Anti-Estrogen Drugs
Tamoxifen (Nolvadex)	Feb 2003	
Raloxifene (Evista)	Dec 1997	
Fulvestrant (Faslodex)	Apr 2002	
Trastuzumab (Herceptin)	Sep 1998	Targeted Biologics
Bevacizumab (Avastin)	Feb 2004	
Everolimus (Afinitor)	Mar 2009	
Pertuzumab (Perjeta)	Jun 2012	
Ado-trastuzumab emtansine (Kadcyla)	Feb 2013	
Palbociclib (Ibrance)	Feb 2015	
Tykerb (Lapatinib)	Sep 2015	
Zoledronate (Zometa)	Aug 2001	Biphosphonates
Pamidronate (Aredia)	May 2002	
Alendronate (Fosamex)	Feb 2008	
Denosumab (Xgeva)	Jun 2010	
Ibandronate (Boniva)	Apr 2012	
Risedronate (Actonel)	Jun 2014	
Doxorubicin (Adriamycin)	Dec 1987	Anthracyclines
Mitoxantrone (Novantrone)	Apr 2006	
Epirubicin (Ellence)	Sep 2008	
Liposomal Doxorubicin (Doxil)	Feb 2013	
Anastrozole (Arimidex)	Jun 2010	Aromatase Inhibitors
Exemestane (Aromasin)	Apr 2011	
Letrozole (Femara)	Jun 2011	
Docetaxel & Cyclophosphamide	N/A	Combination Chemotherapy
Docetaxel, Doxorubicin & Cyclophosphamide	N/A	
Docetaxel & Carboplatin	N/A	
Paclitaxel & Capecitabine	N/A	
Docetaxel & Capecitabine	N/A	
Docetaxel & Carboplatin	N/A	
Paclitaxel & Carboplatin	N/A	
Paclitaxel & Capecitabine	N/A	
Paclitaxel & Carboplatin	N/A	
Irinotecan & Temozolomide	N/A	
Gemcitabine & Carboplatin	N/A	
Ixabepilone & Capecitabine	N/A	
Doxorubicin & Cyclophosphamide	N/A	
Doxorubicin, Cyclophosphamide & Paclitaxel	N/A	
Doxorubicin, Cyclophosphamide & Docetaxel	N/A	
Epirubicin & Cyclophosphamide	N/A	
Cyclophosphamide, Doxorubicin, & Fluorouracil	N/A	
Cyclophosphamide, Methotrexate & 5-Flourouracil	N/A	
5-Flourouracil, Doxorubicin & Cyclophosphamide	N/A	
5-Flourouracil, Epirubicin & Cyclophosphamide	N/A	

Sources: National Cancer Institute (2016), U.S. Food and Drug Administration (2017c)

Table B4 FDA-approved HIV drugs.

Drug (Brand Name)	Approval	Drug Class
Zidovudine (Retrovir)	Mar 1987	Nucleoside
Didanosine (Videx)	Oct 1991	Reverse
Stavudine (Zerit)	Jun 1994	Transcriptase
Lamivudine (Epivir)	Nov 1995	Inhibitors
Abacavir (Ziagen)	Dec 1998	(NRTIs)
Didanosine (Videx EC)	Oct 2000	
Tenofovir Disoproxil Fumarate (Viread)	Oct 2001	
Emtricitabine (Emtriva)	Jul 2003	
Saquinavir (Invirase)	Dec 1995	Protease
Idinavir (Crixivan)	Mar 1996	Inhibitors
Ritonavir (Norvir)	Mar 1996	
Nelfinavir (Viracept)	Mar 1997	
Atazanavir (Reyataz)	Jun 2003	
Fosamprenavir (Lexiva)	Oct 2003	
Tipranavir (Aptivus)	Jun 2005	
Darunavir (Prezista)	Jun 2006	
Nevirapine (Viramune)	Jun 1996	Non-Nucleoside
Delavirdine (Rescriptor)	Apr 1997	Reverse
Efavirenz (Sustiva)	Sep 1998	Transcriptase
Etravirine (Intelence)	Jan 2008	Inhibitors
Nevirapine (Viramune XR)	Mar 2011	(NNRTIs)
Rilpivirine (Edurant)	May 2011	
Lamivudine & Zidovudine (Combivir)	Sep 1997	Combination
Lopinavir & Ritonavir (Kaletra)	Sep 2000	Medications
Abacavir, Lamivudine & Zidovudine (Trizivir)	Nov 2000	
Abacavir & Lamivudine (Epzicom)	Aug 2004	
Emtricitabine & Tenofovir Disoproxil Fumarate (Truvada)	Aug 2004	
Efavirenz, Emtricitabine & Tenofovir Disoproxil Fumarate (Atripla)	Jul 2006	
Emtricitabine, Rilpivirine & Tenofovir Disoproxil Fumarate (Complera)	Aug 2011	
Cobicistat, Elvitegravir, Emtricitabine & Tenofovir Disoproxil Fumarate (Stribild)	Aug 2012	
Abacavir, Dolutegravir & Lamivudine (Triumeq)	Aug 2014	
Atazanavir & Cobicistat (Evotaz)	Jan 2015	
Cobicistat & Darunavir (Prezcobix)	Jan 2015	
Cobicistat, Elvitegravir, Emtricitabine & Tenofovir Alafenamide Fumarate (Genvoya)	Nov 2015	
Emtricitabine, Rilpivirine & Tenofovir Alafenamide Fumarate (Odefsey)	Mar 2016	
Enfuvirtide (Fuzeon)	Mar 2003	Fusion Inhibitors
Maraviroc (Selzentry)	Aug 2007	Entry Inhibitors
Raltegravir (Isentress)	Oct 2007	Integrase
Dolutegravir (Tivicay)	Aug 2013	Inhibitors
Elvitegravir (Vitekta)	Sep 2014	
Cobicistat (Tybost)	Sep 2014	Pharmacokinetic Enhancers

Sources: AidsInfo (2017), U.S. Food and Drug Administration (2016a, 2017c)

Table B5 FDA-approved hypertension drugs.

Drug (Brand Name)	Approval	Drug Class
Reserpine (Raudixin)	Mar 1955	Antiadrenergic
Guanadrel (Hylorel)	Dec 1982	
Methyldopa (Aldomet)	Feb 1986	
Clonidine (Catapres)	Jul 1987	
Prazosin (Minipress)	Sep 1988	
Guanabenz	Apr 1995	
Phentolamine (Regitine)	Mar 1998	
Terazosin (Hytrin)	Mar 1998	
Doxazosin (Cardura)	Oct 2000	
Guanfacine (Tenex)	Oct 2012	
Phenoxybenzamine (Dibenzyline)	Jan 2017	
Guanethidine (Ismelin)	N/A	
Deserpidine (Harmonyl)	Apr 1957	Angiotensin Converting
Captopril (Capoten)	Feb 1996	Enzyme (ACE) Inhibitor
Enalapril (Vasotec)	Jan 2001	
Lisinopril (Prinivil)	Jul 2002	
Moexipril (Univasc)	May 2003	
Benazepril (Lotensin)	Feb 2004	
Fosinopril (Monopril)	May 2005	
Quinapril (Accupril)	Jun 2006	
Trandolapril (Mavik)	Jun 2007	
Ramipril (Altace)	Jun 2008	
Perindopril (Coversyl)	Nov 2009	
Amlodipine & Perindopril (Prestalia)	Jan 2015	
Chlorothiazide (Diuril)	Sep 1958	Diuretics
Polythiazide (Renese)	Sep 1961	
Hydrochlorothiazide (Microzide)	Jan 1973	
Furosemide (Lasix)	Oct 1981	
Methyclothiazide	Jun 1982	
Hydroflumethiazide (Saluron)	May 1985	
Amiloride (Midamor)	Jan 1986	
Spironolactone (Aldactone)	Jul 1986	
Triamterene-Hydrochlorothiazide (Dyazide)	Dec 1987	
Atenolol-Chlorthalidone (Tenoretic)	Jul 1992	
Indapamide (Lozol)	Jul 1995	
Bumetanide (Bumex)	Nov 1996	
Metolazone (Zaroxolyn)	Dec 2003	
Torsemide (Demadex)	May 2005	
Ethacrynic Acid (Edecrin)	Jul 2015	
Deserpidine-Methyclothiazide (Enduronyl)	Aug 1961	Combination Therapy
Reserpine-Polythiazide (Renese-R)	Oct 1963	
Reserpine-Chlorthalidone (Regroton)	May 1964	
Reserpine-Methyclothiazide (Diutensen-R)	Sep 1975	
Reserpine-Hydrochlorothiazide (Hydroserpine)	Jan 1977	
Hydralazine-Reserpine-Hydrochlorothiazide (Hydrap-ES)	Sep 1977	
Hydralazine-Hydrochlorothiazide (Apresazide)	Sep 1977	
Timolol-Hydrochlorothiazide (Timolide)	Dec 1981	
Reserpine-Chlorothiazide (Diupres)	May 1982	
Reserpine-Hydroflumethiazide	Mar 1983	
Reserpine-Trichlormethiazide	Apr 1983	
Methyldopa-Hydrochlorothiazide (Aldoril)	Feb 1987	
Propranolol-Hydrochlorothiazide (Inderide)	Apr 1987	
Spironolactone-Hydrochlorothiazide (Aldactazide)	Jul 1987	
Triamterene-Hydrochlorothiazide (Dyazide)	Dec 1987	
Clonidine-Chlorthalidone (Combipres)	Dec 1987	
Amiloride Hydrochlorothiazide (Moduretic)	May 1988	
Atenolol-Chlorthalidone (Tenoretic)	Jul 1992	
Enalapril-Diltiazem (Teczem)	Oct 1996	
Enalapril Felodipine (Lexxel)	Dec 1996	
Captopril-Hydrochlorothiazide (Capozide)	Dec 1997	
Bisoprolol-Hydrochlorothiazide (Ziac)	Sep 2000	
Enalapril-Hydrochlorothiazide (Vaseretic)	Sep 2001	
Eprosartan-Hydrochlorothiazide (Teveten HCT)	Nov 2001	
Lisinopril-Hydrochlorothiazide (Zestoretic)	Jul 2002	
Benazepril-Hydrochlorothiazide (Lotensin HCT)	Feb 2004	
Metoprolol-Hydrochlorothiazide (Lopressor HCT)	Aug 2004	
Moexipril-Hydrochlorothiazide (Uniretic)	Mar 2007	
Nadolol-Bendroflumethiazide (Corzide)	Mar 2007	
Amlodipine-Benazepril (Lotrel)	May 2007	
Quinapril-Hydrochlorothiazide (Accuretic)	Aug 2007	
Aliskiren-Valsartan (Valturna)	Sep 2009	

Table B5 FDA-approved hypertension drugs (continued).

Drug (Brand Name)	Approval	Drug Class
Losartan-Hydrochlorothiazide (Hyzaar)	Oct 2010	Combination Therapy
Aliskiren-Hydrochlorothiazide (Amturnide)	Dec 2010	(continued)
Telmisartan-Hydrochlorothiazide (Micardis)	Sep 2011	
Irbesartan-Hydrochlorothiazide (Avalide)	Sep 2012	
Valsartan-Hydrochlorothiazide (Diovan)	Sep 2012	
Candesartan-Hydrochlorothiazide (Atacand)	Dec 2012	
Amlodipine-Valsartan (Exforge)	Mar 2013	
Amlodipine-Atorvastatin (Caduet)	Nov 2013	
Amlodipine-Telmisartan (Twynsta)	Jan 2014	
Amlodipine-Valsartan-Hydrochlorothiazide (Exforge HCT)	Jun 2015	
Olmesartan-Hydrochlorothiazide (Benicar HCT)	Oct 2016	
Amlodipine-Olmesartan (Azor)	Nov 2016	
Deserpidine-Hydrochlorothiazide	N/A	
Guanethidine-Hydrochlorothiazide (Esimil)	N/A	
Methyldopa-Chlorothiazide (Aldoclor)	N/A	
Hydralazine (Apresoline)	Oct 1978	Vasodilators
Minoxidil	Jul 1999	
Mecamylamine (Inversine)	Mar 2013	
Propranolol (Inderal)	Nov 1985	Beta Blockers
Penbutolol (Levatol)	Dec 1987	
Atenolol (Tenormin)	Jan 1992	
Nadolol (Corgard)	Oct 1993	
Metoprolol (Lopressor)	Dec 1993	
Pindolol (Visken)	Jan 1994	
Acebutolol (Sectral)	Apr 1995	
Timolol (Betimol)	Mar 1997	
Labetalol (Trandate)	Aug 1998	
Betaxolol (Kerlone)	Oct 1999	
Carteolol (Ocupress)	Jan 2000	
Bisoprolol (Zebeta)	Jun 2001	
Esmolol (Brevibloc)	May 2005	
Carvedilol (Coreg)	Sep 2007	
Nebivolol (Bystolic)	Jul 2015	
Penbuterol	N/A	
Verapamil (Calan)	Jul 1992	Calcium Channel Blockers
Nicardipine (Cardene)	Dec 1996	
Diltiazem (Cardizem)	Dec 1999	
Isradipine (DynaCirc)	Apr 2006	
Amlodipine (Norvasc)	Jun 2007	
Felodipine (Plendil)	Apr 2008	
Nifedipine (Procardia)	Jun 2010	
Nisoldipine (Sular)	Jan 2011	
Aliskiren (Tekturna)	Mar 2007	Other Renin-Angiotensin
Eplerenone (Inspra)	Aug 2008	System Antagonists
Losartan (Cozaar)	Oct 2010	Angiotensin II Receptor
Eprosartan (Teveten)	Nov 2011	Blockers
Azilsartan and Chlorthalidone (Edarbyclor)	Dec 2011	
Irbesartan (Avapro)	Oct 2012	
Candesartan (Atacand)	Jan 2014	
Telmisartan (Micardis)	Jul 2014	
Valsartan (Diovan)	Jun 2015	
Nevivolol and Valsartan (Byvalson)	Jun 2016	
Amlodipine and Olmesartan (Olmesartan)	Oct 2016	

Sources: U.S. Food and Drug Administration (2017c)

Table B6 List of FDA-approved drugs that were withdrawn from the market.

Disease	Drug	Approval	Withdrawal	Time on Market
Breast cancer	Avastin*	Feb 2004	Nov 2011	7.8 years
HIV	Hivid	Jun 1992	Dec 2006	14.5 years
Hypertension	Ticrynafen	May 1979	Jun 1982	2.7 years
Hypertension	Posicor	Jun 1997	Jun 1998	1.0 year
Hypertension	Valturna	Sep 2009	Jul 2012	2.8 years

* Avastin's indication for breast cancer was removed but the drug itself remained on the market.