Submitted to *Operations Research*
manuscript MS-0001-1922.65

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

The FDA requires clinical trial evidence that is statistically significant at the 2.5% level when approving novel drugs, but the agency often uses regulatory discretion when interpreting these standards. Factors such as target disease severity, prevalence, and availability of existing therapies are qualitatively considered, yet no quantitative framework is used to evaluate how such characteristics impact approval standards. We propose a novel queueing network model to analyze the drug approval process, which explicitly incorporates these factors, as well as obsolescence among drugs. Given an objective of maximizing health benefits plus the monetary value of drug approval and rejection, the optimal policy relaxes approval standards for diseases with lengthy clinical trials, greater attrition rates in the development stage, or low intensity of new drug development. We estimate model parameters and calculate the optimal significance levels for drugs targeting breast cancer, HIV, and hypertension. Our results suggest that a significance level of 2.5% is too stringent for some diseases yet too lenient for others. We perform a counterfactual analysis to evaluate the impact of the FDA Fast Track program and find that this program achieves a level of societal health benefit that cannot be attained by merely changing approval standards.

Key words: FDA, Drug Approval, Queueing Model, Healthcare Policy
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 In September 2004, for example, the anti-inflammatory drug Vioxx was withdrawn from world markets due to safety concerns over increased risks of heart attack and stroke, after more than 160,000 patients suffered adverse events and 38,000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*—a measure of a disease's impact on both mortality (length of life) and morbidity (quality of life) in a patient population, *prevalence*—the number of individuals afflicted with a disease, intensity of research and development (R&D), and the number of alternative treatments available for a target condition. In this paper, we refer to a *drug* as a substance intended to diagnose, cure, treat, or prevent disease; we 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.

Current FDA policy requires pharmaceutical companies to first demonstrate that a 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*. 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 (*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).

FDA guidelines for drug approval recommend a constant threshold of $\alpha = 2.5\%$ for all diseases (FDA 2017e), which has both benefits and challenges. By prioritizing diseases equally and holding all drugs to the same efficacy standards, this policy is impartial. The choice of $\alpha = 2.5\%$ is arbitrary, however, 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 exist. A fixed threshold ignores the nuances of clinical trial design (e.g., rate of new molecule discovery, trial duration, 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).

In recognition of the limitations of a fixed threshold, the FDA has introduced programs that provide the agency with 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 is the Orphan Drug Act of 1983. In an attempt to offset the high costs of drug development, the U.S. Congress established tax credits and market exclusivity rights for companies that create drugs for rare diseases (FDA 2017b). 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 the 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. Raising the significance level (making it easier to approve drugs) for diseases with few drugs in development increases the chance of approving a potentially harmful drug, but for patients with few treatment alternatives, the benefits of permitting more drugs to enter the market may outweigh the costs.

(ii) The FDA also partially considers disease severity. 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 re-approved Lotronex in 2002 with restricted use (FDA 2016a). The FDA's consideration of disease severity is indicated in 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 (FDA 2015). The *Fast Track* program facilitates development and 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 heart disease is cholesterol level. A *Breakthrough Therapy* designation aims to hasten the development and review of drugs that demonstrate a significant clinical improvement over existing therapies. Finally, *Priority Review* requires 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: vary the FDA's choice of significance level based on characteristics of 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 (FDA 2017d). By developing

a model in which the significance level required for approval explicitly depends on the characteristics of the drug development process, one can discern the quantitative effect of a given factor on the likelihood of approval. Furthermore, the FDA is often accused of fostering opaque approval policies, and a more objective approach to drug approval could improve the transparency of this process.

The contributions of this paper are as follows:

- We develop a queueing approach to analyze the drug development and approval process, accounting for characteristics such as disease severity and prevalence, R&D intensity, clinical trial duration, and the availability of alternative treatments. We model the drug development process using a series of $M/M/\infty$ queues, and the market for approved drugs as a collection of $M/M/1/1$ and $M/M/\infty$ queues. 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 societal expected net benefits from approving and rejecting drugs, which include the health impact of drugs on the market, as well as monetary values for the correct decision of approving effective drugs and for the incorrect decisions of approving ineffective (type I error) and rejecting effective (type II error) drugs. We interpret health impacts as the incremental gain in QALYs associated with novel drugs and monetary values as the change in the market capitalization of publicly traded pharmaceutical companies following news of successful drug approval, rejection, or withdrawal. We show that the optimal significance level is higher (easier to approve) 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 and post-approval market. 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.
- We perform a counterfactual analysis to evaluate the effect of the Fast Track program on health impacts and monetary values. Using published studies on the observed effects of Fast Track on clinical trial duration and NDA review, we estimate parameters for a hypothetical approval process without this program. Our results indicate that, by bringing drugs to the market more quickly, Fast Track increases both health benefits and societal monetary value. Furthermore, we find that Fast Track attains a level of health benefit that cannot be achieved by only changing the significance level for drug approval.

2. Related Literature

Drug Development and Approval. Three oft-cited sources of inefficiency in the current drug approval process 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). Additionally, the FDA has been criticized for 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. Despite the FDA's recommendation that drugs should be tested against an active comparator or placebo in two randomized, double-blind trials, more than 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 or placebo. While this demonstrates the agency's flexibility in considering a wide range of clinical trial evidence, it obfuscates the agency's 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 et al. (2017), who use Bayesian Decision Analysis to quantify how the FDA's approval policy could depend on the burden of disease and patient preferences. The authors compute the optimal statistical significance level for the 23 most common types of cancer and argue that the traditional level of $\alpha = 2.5\%$ is too low for rare cancers with few treatment options and short survival times, and too high for prevalent cancers with many treatment options and long survival times. Our work differs from this study in several aspects. First, Montazerhodjat et al. (2017) assume a static model, setting the significance level based on the current disease characteristics, while our model is dynamic and considers the system-wide effects of changing approval standards. In Montazerhodjat et al. (2017), the choice of significance level is driven by the *current* availability of treatments for a disease, while in our work, we consider what impact changing the significance level will have on *future* treatment availability. Another importance difference is that our work attempts to capture phenomena such as substitution between drugs within the same therapeutic class and obsolescence of older therapies, effects that are ignored in Montazerhodjat et al. (2017).

Randomized Controlled Trials (RCTs). A major bottleneck in the drug approval process is 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 et al. (2017) explicitly incorporate the costs associated with treating patients with a potentially harmful drug and use expected cost analysis to determine the optimal sample size for a balanced two-arm RCT. 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 suggest new combinations of drug regimens. 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. 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 among projects in development. Adler et al. (1995) model the product development process as a queueing network, to 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, but we additionally capture characteristics of the post-approval process, such as obsolescence among drugs. Adler et al. (1995) take the perspective of a single firm, with the objective of maximizing profit, while we take the perspective of the social planner, with the objective of maximizing expected societal benefit.

Other NPD research focuses on the market stage of development, examining questions such as how products compete for market share. 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

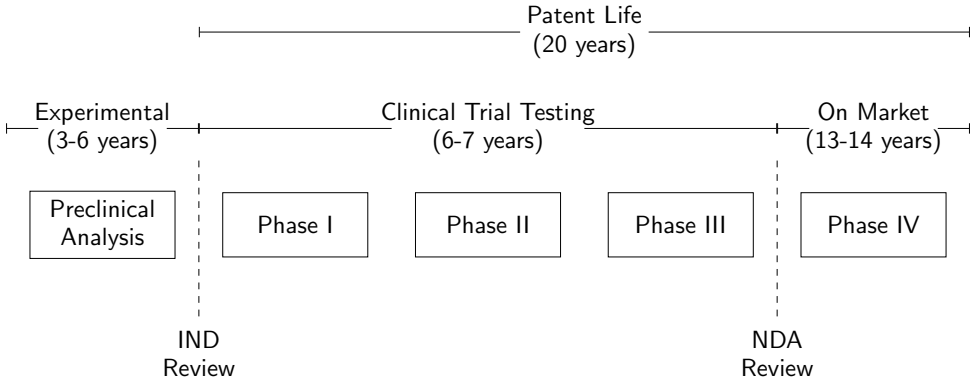
pursued by a firm as a decision variable, whereas R&D intensity is an exogenous parameter in our work. Furthermore, rather than studying market competition for revenue, we examine the role of obsolescence among FDA-approved drugs targeting the same condition.

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 among 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 a potential new pharmaceutical compound and ending with the FDA deciding whether to grant marketing approval to the drug. See Figure 1 for a summary and average duration of each stage (PhRMA 2015a). The creation of 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.

Figure 1 The FDA drug development and approval process.



Note. For each new compound, the FDA reviews two applications submitted by the pharmaceutical company: an IND (Investigational New Drug) and an NDA (New Drug Application).

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 a plan for 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 trials usually consist of three phases, designed to test if the candidate drug is both safe and effective in humans. Phase I involves testing the candidate drug on healthy volunteers to observe potential side effects and the drug's pharmacokinetics (how the drug is metabolized). 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, by comparing patients receiving the candidate drug to those treated with a placebo or standard therapy. The focus of the final stage of clinical testing, Phase III, 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.

At any point during the preclinical or clinical trial stages, the sponsoring pharmaceutical company may choose to withdraw the drug from development. Typical reasons for halting development include the inability to demonstrate efficacy, safety concerns, pharmacokinetic issues, market competition, and financial considerations (Arrowsmith and Miller 2013). After the completion of Phase III trials, the company can submit a New Drug Application (NDA) to the FDA, consisting of clinical trial results as well as a proposal for manufacturing and labeling the medication. The FDA performs a risk-benefit assessment using this information, including data on demonstrated efficacy and reported adverse events, and decides whether the potential benefits of the medication outweigh its risks. Periodically, companies are asked to perform additional testing before being awarded marketing approval (FDA 2014b).

Drugs that ultimately gain FDA approval may then be legally marketed in the U.S and 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, which is usually before initiating clinical trial testing, although applications can be submitted at any point during the development process. Exclusivity, or exclusive marketing rights, are granted by the FDA, with all new drugs receiving 5 years of exclusivity upon approval. 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 (FDA 2016b). Most approved 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 (FDA 2017c).

3.1. Randomized Controlled Trial Design

RCTs 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 RCT, a commonly used design where patients are randomly assigned to a *treatment* group or a *control* 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*—outcomes that represent direct clinical benefit—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.

The main criteria used by the FDA when deciding whether to approve a candidate drug are *safety* and *efficacy*. Safety is assessed by considering the number and type of adverse events experienced by trial volunteers. Efficacy is assessed by monitoring one or more target disease endpoints, 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 3.2) 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 can be objectively measured and must be specified before initiating the trial. 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.

3.2. 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 a standard therapy. We assume that x_i is drawn from a distribution with mean μ_x and variance σ^2, and y_i is drawn from a 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 superiority trials, which assumes that the experimental drug has no effect or a positive effect, compared to the standard therapy. We perform the following hypothesis test, wherein a 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, and ineffective otherwise:

$$H_0 : \delta = 0 \text{ (drug is ineffective)}$$
$$H_1 : \delta > 0 \text{ (drug is effective)}$$

using the following Wald statistic from the observed data:

$$Z_n = (\overline{x} - \overline{y}) \sqrt{I_n}$$

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. By the Central Limit Theorem, Z_n follows a normal distribution with mean $\delta\sqrt{I_n}$ and variance 1. Based on Z_n, one computes the p-value associated with the Wald statistic. If the p-value is less than a threshold 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 H_1). Given an approval policy α and prior p, we obtain the following joint probability expressions for approved effective (AE), approved ineffective (AI), rejected effective (RE), and rejected ineffective (RI) drugs:

$$\begin{aligned}
\pi_{\text{AE}}(\alpha) &= [1 - \Phi(\Phi^{-1}(1-\alpha) - \delta\sqrt{I_n})]\, p && \text{(Approving an effective drug)} \\
\pi_{\text{AI}}(\alpha) &= \alpha\,(1-p) && \text{(Approving an ineffective drug)} \\
\pi_{\text{RE}}(\alpha) &= \Phi(\Phi^{-1}(1-\alpha) - \delta\sqrt{I_n})\, p && \text{(Rejecting an effective drug)} \\
\pi_{\text{RI}}(\alpha) &= (1-\alpha)\,(1-p) && \text{(Rejecting an ineffective drug)}
\end{aligned} \tag{1}$$

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

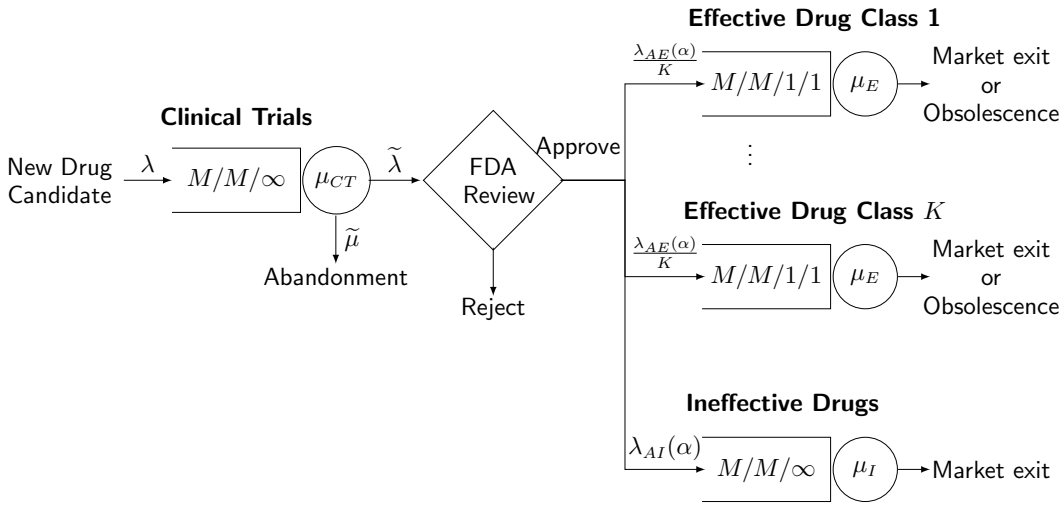
Candidate drugs begin clinical trials according to a Poisson process with rate λ. Drugs either complete clinical trials successfully, or the sponsoring company halts the trials early. We assume that the time until a drug company abandons the trial is exponentially distributed with rate μ_{AB}, and the time to complete trials is exponentially distributed with rate μ_{CT}. Drugs in clinical trials are considered in "service"; following service completion, drugs advance to FDA review with probability $\frac{\mu_{CT}}{\mu_{CT}+\mu_{AB}}$ or 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 trial abandonment rate as $\tilde{\mu} = \lambda \frac{\mu_{AB}}{\mu_{CT}+\mu_{AB}}$.

After FDA review, drugs gain approval if the p-value associated with the clinical trial demonstrating efficacy is less than the chosen significance level α, and are denied approval otherwise. For simplicity, we assume that the FDA's decision occurs immediately, though in reality the review process takes between six months and two years. This delay could be accounted for by modeling the review stage as an $M/M/\infty$ queue, but would not change our results. In steady state, the output of the FDA review stage constitutes a thinning of a Poisson process into four separate and independent Poisson processes. Using the probabilities in (1), the arrival rates for each of these Poisson processes are:

$$\lambda_{AE}(\alpha) = \tilde{\lambda}\pi_{AE}(\alpha), \quad \lambda_{AI}(\alpha) = \tilde{\lambda}\pi_{AI}(\alpha), \quad \lambda_{RE}(\alpha) = \tilde{\lambda}\pi_{RE}(\alpha), \quad \lambda_{RI}(\alpha) = \tilde{\lambda}\pi_{RI}(\alpha). \tag{2}$$

After undergoing FDA review, rejected drugs depart the system, while approved drugs enter the market. Approved ineffective drugs and approved effective drugs differ with respect to how long they spend on the market. Approved *ineffective* drugs spend relatively little time on the market as

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



Note. $\tilde{\lambda} = \lambda \frac{\mu_{CT}}{\mu_{CT}+\mu_{AB}}$ and $\tilde{\mu} = \lambda \frac{\mu_{AB}}{\mu_{CT}+\mu_{AB}}$.

we assume that they are discontinued by dissatisfied patients. Approved *effective* drugs typically spend decades on the market and may, eventually, become obsolete as newer drugs enter the market.

Given these differences, we model effective and ineffective FDA-approved drugs separately. Ineffective drugs are modeled using an $M/M/\infty$ queue, where "service" represents time on the market $1/\mu_I$ before withdrawal of the product. The market for effective drugs is modeled using a collection of K parallel $M/M/1/1$ queues with an average service time of $1/\mu_E$. When a new effective drug is approved, it is equally likely to join each of the K queues. If no other drugs are in service, the drug begins service immediately. If an existing drug is in service, then this drug exits the system and the new drug begins service. This preemption 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 $M/M/1/1$ queue with a drug class and view K as the number of unique drug classes available to treat a particular disease.

In our analysis, we consider the system in steady state and assume that the parameters are time invariant. We assume that the FDA's objective is to choose the statistical significance level α to maximize the societal expected net benefit from approving and rejecting drugs, which consists of two components: health impacts, measured in terms of quality-adjusted life years (QALYs), and monetary values, measured in U.S. dollars. By measuring health impacts in QALYs, we account for the effects of a drug on both the length and quality of life of a patient. In accordance with the notion that patient health increases when there are more effective treatments available and decreases when an ineffective drug is prescribed, we associate a health benefit Q_E to each effective drug on the market, and a health cost Q_I to each ineffective drug. Additionally, each time a new drug is approved or rejected, the market gains or loses value according to perceived changes in the profitability of pharmaceutical companies. Accordingly, we let C_{AE} denote the monetary gain associated with approving an effective drug, and let C_{AI} and C_{RE} denote the monetary losses associated with approving ineffective (type I error) and rejecting effective (type II error) drugs,

Table 1 Summary of key model parameters.

Before FDA review		After FDA review	
σ	Standard deviation of the candidate drug response	K	Number of unique drug classes on the market
δ	Treatment effect of a candidate drug	Q_E	Per drug health benefit of an effective drug
p	Prior probability that candidate drug is effective	Q_I	Per drug health cost of an ineffective drug
n	Clinical trial enrollment	C_{AE}	Per drug monetary gain of approving effective drugs
λ	Rate that drugs initiate clinical trials	C_{AI}	Per drug monetary loss of approving ineffective drugs
μ_{CT}	Rate that clinical trials are completed	C_{RE}	Per drug monetary loss of rejecting effective drugs
μ_{AB}	Rate that firms abandon clinical trials	WTP	Willingness to pay per QALY
$\tilde{\lambda}$	Rate that drugs enter FDA review	$1/\mu_E$	Average market life of an effective drug
		$1/\mu_I$	Average market life of an ineffective drug

respectively. The monetary value associated with rejecting an ineffective drug is normalized to zero. To make the health impact and monetary values directly comparable, we multiply QALYs by the willingness to pay (WTP), the maximum amount that an individual would be willing to pay per QALY gained (Drummond et al. 2003).

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

$$\begin{aligned}\alpha^* &= \arg\max_{\alpha\in[0,1]} V(\alpha) \\ &= \arg\max_{\alpha\in[0,1]} \big\{\text{Net health impact}\cdot\text{WTP} + \text{Net monetary value}\big\}. \\ &= \arg\max_{\alpha\in[0,1]} \big\{\big(Q_E\mathbb{E}[N_E(\alpha)] - Q_I\mathbb{E}[N_I(\alpha)]\big)\text{WTP} + (C_{AE}\lambda_{AE}(\alpha) - C_{AI}\lambda_{AI}(\alpha) - C_{RE}\lambda_{RE}(\alpha))\big\}.\end{aligned} \tag{3}$$

The per drug health benefit or cost is multiplied by the expected number of effective or ineffective drugs, $\mathbb{E}[N_E(\alpha)]$ or $\mathbb{E}[N_I(\alpha)]$, respectively. Letting $\psi_E(\alpha) = \lambda_{AE}(\alpha)/(K\mu_E)$ and $\psi_I(\alpha) = \lambda_{AI}(\alpha)/\mu_I$, we can write these terms as:

$$\mathbb{E}[N_E(\alpha)] = \frac{K\psi_E(\alpha)}{1+\psi_E(\alpha)}, \qquad \mathbb{E}[N_I(\alpha)] = \psi_I(\alpha). \tag{4}$$

Each monetary value is multiplied by the corresponding rate at which drugs are approved or rejected, which reflects the societal cost (or benefit) of a new drug. Note that this is a one time gain/loss in monetary value.

4.2. Model Analysis

We first examine the structure of the optimal approval policy 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.

The following result shows that the optimal significance level α^* is unique and is the solution to a non-linear equation.

THEOREM 1. *The expected net benefit function $V(\alpha)$ is concave in α, and the optimal policy α^* satisfies the following first order condition:*

$$\alpha^* = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}}\log\left(\frac{1-p}{p}\frac{C_{AI} + WTP\cdot Q_I/\mu_I}{WTP\cdot Q_E/(\mu_E(1+\psi_E(\alpha^*))^2) + C_{RE} + C_{AE}}\right) + \frac{\delta\sqrt{I_n}}{2}\right). \tag{5}$$

Theorem 1 demonstrates that the optimal FDA approval policy, α^*, weighs the steady-state monetary losses and health costs of approving ineffective drugs against the monetary gains (or losses) and health benefits of approving (or rejecting) effective drugs. Although no closed form expression for the optimal policy exists, we can analyze the comparative statics of α^* using the first order condition.

PROPOSITION 1. *The optimal approval policy α^* is*

(a) increasing in Q_E, C_{AE}, C_{RE}, μ_I, and μ_{AB},

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

(c) increasing in p and decreasing in μ_E under the additional assumption that $\psi_E(\alpha^*) < 1$.

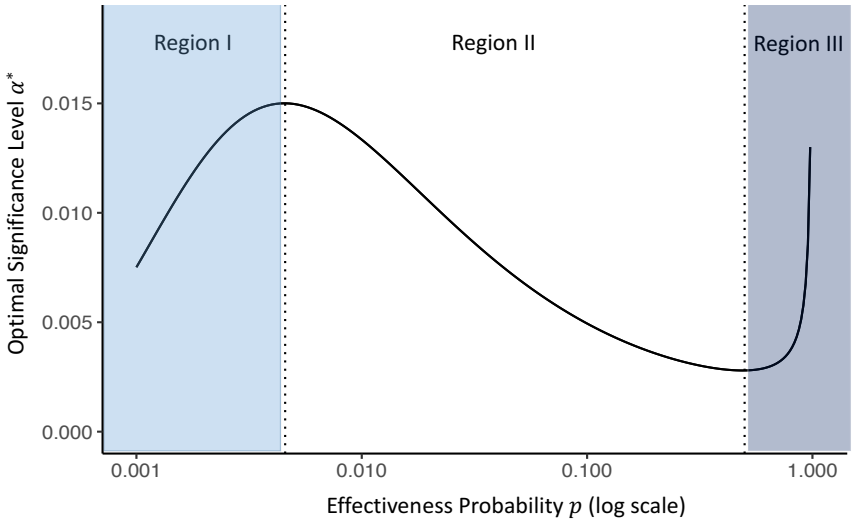
From Proposition 1, we see that it should be more difficult to approve drugs for diseases with many compounds in development (high λ) and short clinical trial durations (large μ_{CT}), and easier to approve drugs for diseases with high attrition rates (large μ_{AB}). Drugs with greater health benefits Q_E (due to their ability to increase length or quality of life) or higher monetary rejection costs C_{RE} (due to a type II error) have less stringent approval policies compared to drugs with higher monetary approval costs C_{AI} (due to a type I error). Prolonging the average time that ineffective drugs spend on the market $1/\mu_I$ exposes patients to ineffective drugs for longer, and thus disincentivizes approval.

As the prior probability p that drugs are effective increases, or as the average time that effective drugs spend on the market $1/\mu_E$ increases, it is natural to expect that the optimal response would be to approve more drugs. However, Proposition 1 implies that this intuition only holds if the rate $\psi_E(\alpha^*) = \lambda_{AE}(\alpha^*)/K$ at which effective drugs in a given class are approved (see Figure 2) is constrained to be less than one. Because we model the market for effective drugs as a collection of $M/M/1/1$ queues, $\psi_E(\alpha^*) < 1$ is not needed for stability; however we see that this condition serves to limit the degree of crowding in the market. To understand how market crowding results in non-monotonicity of the optimal policy, consider the following example, illustrated in Figure 3.

Consider a disease with a high rate of R&D intensity $\tilde{\lambda}$, and a high health benefit associated with effective drugs Q_E relative to the health cost of ineffective drugs Q_I. For simplicity, suppose that there is no monetary value associated with approving or rejecting drugs, i.e. $C_{AE} = C_{AI} = C_{RE} = 0$. We explain the non-monotonic behavior of the optimal approval policy by considering two key characteristics: the effectiveness probability p and the degree of crowding in the market for approved effective drugs, $\mathbb{E}[N_E(\alpha)]$. In this example, let's define drugs with a low effectiveness probability ($p < 0.5$) as *long shots*, and those with high effectiveness probability ($p \geq 0.5$) as *safe bets*. We consider the market for approved effective drugs *crowded* if many therapies are available ($\mathbb{E}[N_E(\alpha)] \approx K$) or *neglected* if few are available ($\mathbb{E}[N_E(\alpha)] << K$).

We divide Figure 3 into three regions. Region I corresponds to diseases with neglected markets and long shot drugs seeking FDA approval. As the probability of effectiveness increases, the optimal policy is to approve more drugs (despite their low effectiveness probability) because of the large health benefits of effective drugs and the paucity of drugs available to patients. In Region II, drugs are still long shots, but the market is more crowded, so the optimal policy is to approve

Figure 3 Example of the sensitivity of the optimal significance level α^* with respect to the effectiveness probability p if Proposition 1c is not satisfied.



Note. $\sigma = 1$, $\delta = 0.10$, $n = 500$, $\tilde{\lambda} = 8$, $K = 1$, $WTP = 1$, $Q_E = 1$, $Q_I = 0.1$, $\mu_E = 0.01$, $\mu_I = 0.10$, $C_{AE} = 0$, $C_{AI} = 0$, and $C_{RE} = 0$. Region I corresponds to $0 \leq p \leq 0.005$, Region II to $0.005 < p \leq 0.5$, and Region III to $0.5 < p \leq 1$.

fewer drugs as the effectiveness probability increases. This occurs because of drug obsolescence, which limits the number of drug classes available on the market, resulting in diminishing marginal health benefits of approving drugs. As the market becomes crowded, the marginal health benefits of approving effective drugs is outweighed by the marginal health costs of approving ineffective drugs; therefore, the optimal policy is to approve fewer drugs. Finally, in Region III, although the market is crowded and additional drugs have diminishing marginal health benefits, the candidate drugs are reasonably safe bets, so each new drug approval generates a positive expected net health benefit. Therefore, the optimal policy in this region is to approve more drugs as the effectiveness probability increases.

Our analysis thus far assumes a fixed number of unique drug classes K available to treat a particular disease. We next examine how varying K affects the optimal policy. One can interpret an increase in the number of drug classes K as corresponding to the approval of a *first in class* drug—a therapy that uses a new and unique mechanism of action for disease treatment. By treating a condition in a novel manner, first in class drugs potentially offer patients a more tolerable set of side effects or serve a patient population for whom current treatments are inadequate.

In the following proposition, let α_j^* denote the optimal policy when there are j drug classes on the market.

PROPOSITION 2. *The optimal approval policies satisfy*

$$\alpha_0^* \leq \alpha_1^* \leq \cdots \leq \alpha_K^* \leq \cdots \leq \alpha_\infty^*$$

where

$$\alpha_0^* = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}} \log\left(\frac{1-p}{p}\frac{C_{AI} + WTP \cdot Q_I/\mu_I}{C_{RE} + C_{AE}}\right) + \frac{\delta\sqrt{I_n}}{2}\right) \tag{6}$$

and

$$\alpha_\infty^* = 1 - \Phi\left(\frac{1}{\delta\sqrt{I_n}} \log\left(\frac{1-p}{p}\frac{C_{AI} + WTP \cdot Q_I/\mu_I}{WTP \cdot Q_E/\mu_E + C_{RE} + C_{AE}}\right) + \frac{\delta\sqrt{I_n}}{2}\right). \tag{7}$$

Proposition 2 states that the optimal approval policy is non-decreasing in the number of drug classes K, an intuitive result. As K increases, more opportunities exist for different therapy classes and thus the optimal policy is to ease approval standards to fill the market. While α_0^* is purely a mathematical lower bound and does not have a direct interpretation in our model, the optimal policy α_1^* might represent a disease with limited treatment options, such as Ebola infection or muscular dystrophy. The upper bound α_∞^*, derived by letting the number of drug classes K approach infinity, represents the optimal policy for a condition such as mild pain, for which a variety of therapies are available.

Changing the number of drug classes on the market affects not only the optimal policy, but also the expected net benefit from approving and rejecting drugs. In the following proposition, let V_j^* denote the optimal expected net benefit when there are j drug classes on the market.

PROPOSITION 3. *The optimal expected net benefit functions satisfy*

$$V_0^* \leq V_1^* \leq \cdots \leq V_K^* \leq \cdots \leq V_\infty^*,$$

and, for all $K \geq 1$, and for any α,

$$V_{K+1}(\alpha) - V_K(\alpha) > V_{K+2}(\alpha) - V_{K+1}(\alpha).$$

The first result in Proposition 3 shows that, intuitively, increasing the number of drug classes K results in higher expected net benefit due to additional health benefits generated by effective drugs on the market. We also find that when few drug classes are available (K is low), an increase in K has a larger effect on expected net benefits compared to when there are many drug classes (K is high). This suggests that innovation in drug development is particularly important for diseases with a dearth of available treatments.

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 and compare them to a traditional policy of $\alpha = 2.5\%$. The goals of this analysis are (i) to understand the impact of the development parameters on the optimal policy, and (ii) to demonstrate how our model can be used to gain insights about disease-specific drug approval recommendations.

5.1. Parameter Estimation

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

Clinical trial parameters. The pre-FDA review parameters are numerically estimated for each disease using clinical trial data from `clinicaltrials.gov`, historical drug approval data from Drugs@FDA, and NDA approval probabilities from Thomas et al. (2016). We estimate the clinical trial completion rate μ_{CT} using the mean durations of Phase I-III trials. We estimate the probability that a drug completes all three phases of clinical trials $\mathbb{P}(\text{Complete clinical trials})$ and estimate the clinical trial abandonment rate as $\mu_{AB} = \frac{\mu_{CT}[1-\mathbb{P}(\text{Complete clinical trials})]}{\mathbb{P}(\text{Complete clinical trials})}$.

We estimate the NDA submission rate $\tilde{\lambda}$ using the average rate of drug approval for a disease (computed using exhaustive lists of approved drugs provided in Appendix Tables B2-B4) and estimates for the NDA approval probability from Thomas et al. (2016). The clinical trial initiation rate λ is estimated using $\tilde{\lambda}$ and $\mathbb{P}(\text{Complete clinical trials})$.

The clinical trial information $\delta\sqrt{I_n}$ is estimated by assuming that the statistical power of the trial—the probability of approving a drug conditional on the drug being effective—is 90%, given a traditional statistical significance level of $\alpha = 2.5\%$. To estimate the probability p that a drug is effective, we select the value of p that makes the probability of approving a drug in our model equal to the NDA approval probability estimates in Thomas et al. (2016), assuming $\alpha = 2.5\%$.

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. Lists of all drug classes and references are provided in Appendix Table B1.

Health impacts. We interpret the per-drug health benefits and costs Q_E and Q_I as the change in QALYs associated with one additional effective or ineffective drug on the market, respectively. We calculate Q_E as the incremental per-drug per-person gain in QALYs associated with newly approved drugs (estimated by Chambers et al. (2017)) multiplied by the new drug's expected

market size. We assume that all patients with a disease are equally likely to take any of the K drug classes available, and that each patient will be treated with one drug. Given these assumptions, we estimate the market size as either the incidence (for acute diseases) or the prevalence (for chronic diseases) of the disease being treated, divided by the number of drug classes K, so that each drug has the same market share.

To calculate Q_I, we assume that the total health cost Q_I/μ_I is proportional to the total health benefit Q_E/μ_E. We use the ratio C_{AI}/C_{AE} of the monetary losses of approving ineffective drugs to the monetary gains of approving effective drugs as our constant of proportionality, with the idea that the relative stock market reactions of approving and withdrawing a drug may also reflect the relationship between health benefits of effective and costs of ineffective drugs.

Monetary values. We define the monetary gains and losses C_{AE}, C_{AI}, and C_{RE} as the average change in market capitalization of pharmaceutical firms in response to the approval of an effective drug, approval of an ineffective drug, and rejection of an effective drug, respectively. We use published estimates of percent abnormal market returns at the time of initial review, the time a drug is announced as approvable, the approval announcement day (or rejection announcement day), the day after the approval announcement, and following market withdrawal (Sarkar and de Jong 2006, Ahmed et al. 2002). We estimate monetary values by combining these published estimates with the market capitalization of pharmaceutical companies to reflect the aggregate monetary gain or loss associated with a drug approval or rejection decision by the FDA. Note that this gain or loss is incurred once for each drug that is approved or rejected.

Market durations. We estimate the average time that effective drugs spend on the market $1/\mu_E$ as the sum of the time that a drug spends on the market on patent $1/\mu_{PAT}$ and the time on the market as a generic or off-patent drug $1/\mu_{GEN}$. To calculate $1/\mu_{PAT}$, we assume that drug companies file patents at the start of preclinical analysis (an average of 4.5 years before Phase I trials), and we subtract the time in preclinical work and in clinical trials from the 20 year standard patent life (PhRMA 2015a). To calculate $1/\mu_{GEN}$, we examine FDA records of drugs (novel and generic) that were discontinued for reasons not related to safety or efficacy between the years of 2015 and 2017 (FDA 2017a).

To estimate the average time that ineffective drugs spend on the market $1/\mu_I$, we calculate the average length of time that withdrawn drugs spend on the market before being removed, for each disease considered. The list of withdrawn drugs and time on the market was obtained from Drugs@FDA and is included in Appendix Table B5.

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

We conduct a numerical study of three high-burden diseases, which 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 2, with additional details provided in Appendix B

Breast cancer is an acute disease responsible for more than 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 suppress the amount of virus in the body, slow disease progression, and substantially prolong life. Currently 1.2 million people in the U.S. 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 often take medications to control their blood pressure throughout their life. Hypertension currently affects 106 million people in the U.S. 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).

Table 2 Parameter estimates for selected diseases.

Parameter	Breast Cancer	HIV	Hypertension	Source
λ (drugs/year)	9.99	4.80	3.85	clinicaltrials.gov, BIO
μ_{CT} (drugs/year)	0.08	0.14	0.31	clinicaltrials.gov
μ_{AB} (drugs/year)	0.46	0.28	0.20	clinicaltrials.gov
$\tilde{\lambda}$ (drugs/year)	1.48	1.60	2.34	BIO
p	0.912	0.985	0.933	BIO
K (classes)	10	6	9	See Appendix Table B1
C_{AE} (billion \$)	0.094	0.094	0.094	Ahmed et al. (2002), Sarkar and de Jong (2006)
C_{AI} (billion \$)	0.102	0.102	0.102	Ahmed et al. (2002), Sarkar and de Jong (2006)
C_{RE} (billion \$)	0.023	0.025	0.024	Ahmed et al. (2002), Sarkar and de Jong (2006)
Q_E (QALYs)	2,350	12,650	1,766,670	CDC.gov, Chambers (2017), NCI
Q_I (QALYs)	7,579	23,986	21,975,400	CDC.gov, Chambers (2017), NCI
WTP (\$/QALY)	100,000	100,000	100,000	Neumann et al. (2014)
μ_E (drugs/year)	0.043	0.039	0.036	FDA.gov, Drugs@FDA
μ_I (drugs/year)	0.128	0.069	0.455	See Appendix Table B6

Note: The clinical trial information $\delta\sqrt{I_n}$ is calculated assuming a 90% statistical power level.
Sources: clinicaltrials.gov (National Library of Medicine and National Institutes of Health 2017); Biotechnology Innovation Organization (BIO) (Thomas et al. 2016); Centers for Disease Control and Prevention (2017); Centers for Disease Control and Prevention (2015); National Cancer Institute (2017); Food and Drug Administration (2017a)

Significant heterogeneity exists in the pre-FDA review timeline across the three diseases (Table 2). 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 1.48 drugs per year. According to Arrowsmith and Miller (2013), this high rate of attrition stems from difficulty in establishing efficacy for oncology drugs in trials with relatively short durations. In contrast, hypertension has a high R&D intensity λ of 3.85 drugs per year, but also the lowest attrition rate, leading to the highest NDA intensity of 2.34 drugs per year. The estimated probability p that a drug is effective, conditional on undergoing FDA review, is similar across the examined conditions, with all estimated probabilities exceeding 0.90.

Substantial variation also exists in the health impact associated with drugs used to treat the three diseases. Hypertension has the highest per-drug health benefit Q_E, while breast cancer medications have the lowest. This effect is driven by both the difference in incremental QALY gain, and the substantially larger market size for hypertension drugs. Effective hypertension drugs also spend more time on the market under patent protection due to their short clinical trial durations, compared to breast cancer drugs which require longer trials, on average. Historically, ineffective hypertension drugs spend the shortest amount of time on the market, potentially because blood pressure is more easily monitored, leading to faster public awareness of a drug's ineffectiveness. The first case of HIV in the United States was identified in 1981, which may partly explain the paucity of drug classes for this disease compared to hypertension and breast cancer, for which treatments have been in development since the 1950s (Department of Health and Human Services 2016).

Using the parameter estimates given in Table 2, we calculate the optimal approval policies α^* for each of the three diseases. Our results, summarized in Table 3, highlight the differences across diseases as well as how each policy compares to a traditional threshold of $\alpha = 2.5\%$. Our model suggests that a stricter approval threshold is optimal for hypertension drugs due to their high NDA intensity and the substantial health costs (due to high prevalence) associated with approving ineffective drugs. In contrast, the optimal approval policy for HIV drugs is less stringent due to their relatively low NDA intensity λ, high effectiveness probability p, and relative paucity of available treatment alternatives.

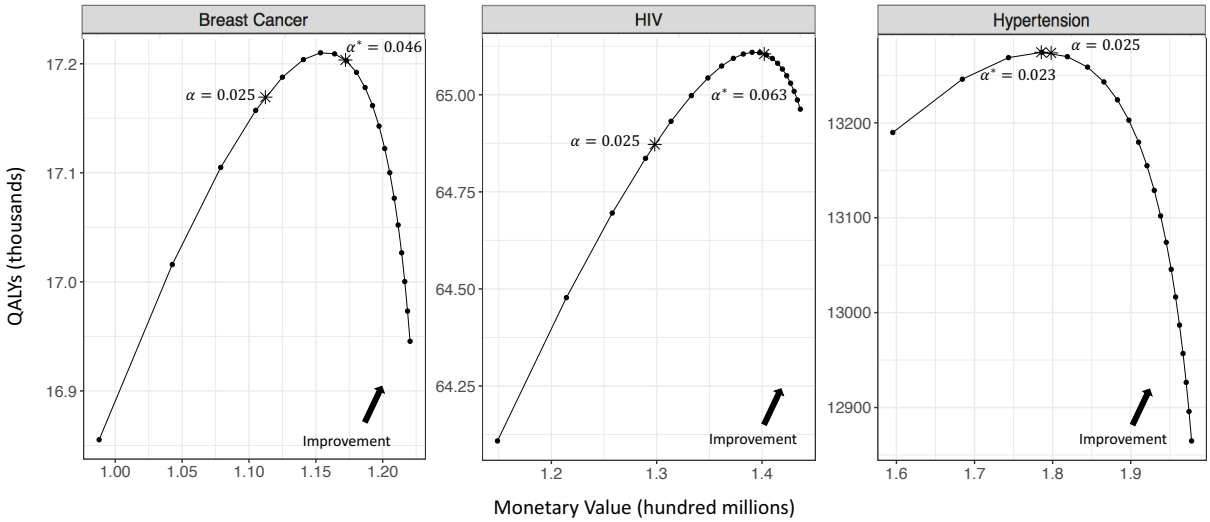
Table 3 Optimal policies for each disease compared to the traditional $\alpha = 2.5\%$ threshold.

	Breast Cancer	HIV	Hypertension
Optimal Policy α^*	0.046	0.063	0.023

Figure 4 depicts the trade-off between monetary value incurred ($C_{AE}\lambda_{AE}(\alpha) - C_{AI}\lambda_{AI}(\alpha) - C_{RE}\lambda_{RE}(\alpha)$) and the number of QALYs achieved ($Q_E\mathbb{E}[N_E(\alpha)] - Q_I\mathbb{E}[N_I(\alpha)]$) for approval policies

ranging from $\alpha = 0.01$ (the furthest left point) to $\alpha = 0.10$ (the furthest right point). In these plots, moving to the upper right represents a desirable policy that increases both monetary value and QALYs. For all three diseases, increasing α from 0.01 to 0.10 always results in higher monetary value because the marginal monetary gains from approving effective drugs outweigh the marginal monetary losses of approving ineffective and rejecting effective drugs. On the other hand, increasing α initially results in more QALYs due to increases in the number of effective drugs available on the market, but eventually leads to fewer QALYs because of diminishing marginal health benefits due to obsolescence. In the cases of breast cancer and HIV, the optimal policy achieves more QALYs and a higher monetary value than a traditional policy, meaning that the optimal policy strictly dominates the traditional policy. On the other hand, the optimal policy for hypertension achieves more QALYs, but a lower net monetary value.

Figure 4 Comparison of the monetary value ($C_{AE}\lambda_{AE}(\alpha) - C_{AI}\lambda_{AI}(\alpha) - C_{RE}\lambda_{RE}(\alpha)$) and QALYs achieved ($Q_E\mathbb{E}[N_E(\alpha)] - Q_I\mathbb{E}[N_I(\alpha)]$) by different approval policies.



Note. Each point on the curve represents a different approval policy α. We vary α from 0.01 (the furthest left point in each plot) to 0.10 (the furthest right point in each plot).

We conduct sensitivity analysis of the optimal approval policies with respect to the nominal parameter values, displayed in Table 2. We focus on three key parameters for our analysis: the prior probability p that a drug is effective, the NDA intensity $\tilde{\lambda}$, and the market duration of effective drugs $1/\mu_E$. The parameter $\tilde{\lambda}$ was chosen because it encompasses the pre-review parameters λ, μ_{CT}, and μ_{AB}. We vary the market duration $1/\mu_E$ and the prior probability p to further explore the non-monotonic relationship between these parameters and the optimal policy.

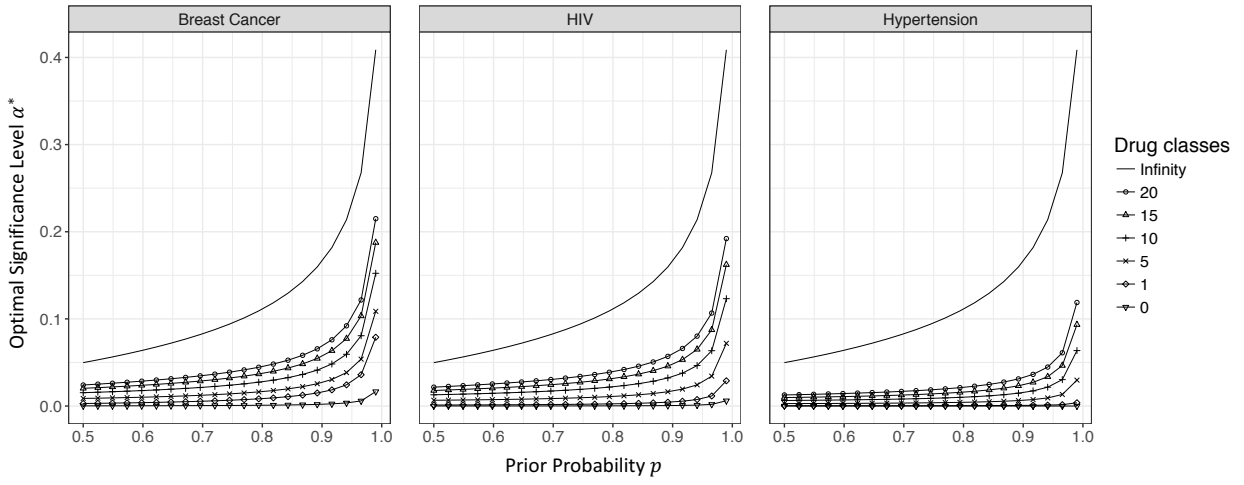
We vary p from 0.5 to 0.99; this range is chosen using the reasoning that the FDA is unlikely to approve a drug if its prior belief that the drug is effective is less than 50%, and it is also unlikely

that the FDA's prior is higher than 99% for any drug. We note that if the probability p is equal to 1, then the optimal policy is to approve all drugs (i.e. $\alpha = 1$). We vary $\tilde{\lambda}$ from 1 drug per year to 4 drugs per year, and $1/\mu_E$ from 5 years to 30 years.

For each parameter, and for each disease, we plot the value of the optimal policy, where the number of drug classes $K = 1, 5, 10, 15$, and 20, along with the lower and upper bounds α_0^* and α_∞^* on the optimal policies (denoted in the plots using $K = 0$ and $K = \infty$, respectively). In the resulting plots, shown in Figures 5 - 7, we see that the optimal approval policies are increasing in the number of drug classes K, as indicated in Proposition 2.

Sensitivity to effectiveness probability. From Figure 5, the optimal policy is increasingly sensitive to changes in the prior probability of drug effectiveness p, as p approaches 1, but is relatively insensitive for $0.5 < p < 0.8$. Our estimates of the efficacy probability are 0.912 (breast cancer), 0.985 (HIV), and 0.933 (hypertension), suggesting that our optimal policies may be sensitive to changes in p, and potentially overestimate overestimate the true optimal α. Figure 5 also highlights the trade-off between the number of drug classes K and effectiveness probability p. For example, with breast cancer, an equivalent optimal policy α^* is obtained if $p = 0.9$ and $K = 10$, or if $p = 0.85$ and $K = 15$. A market with many drug classes (and thus more drug diversity) can withstand a lower probability of effectiveness, to arrive at the same approval policy α^* as a market with relatively fewer unique drug classes but higher effectiveness probability per drug.

Figure 5 Sensitivity of the optimal approval policy to the prior probability p that a drug is effective.



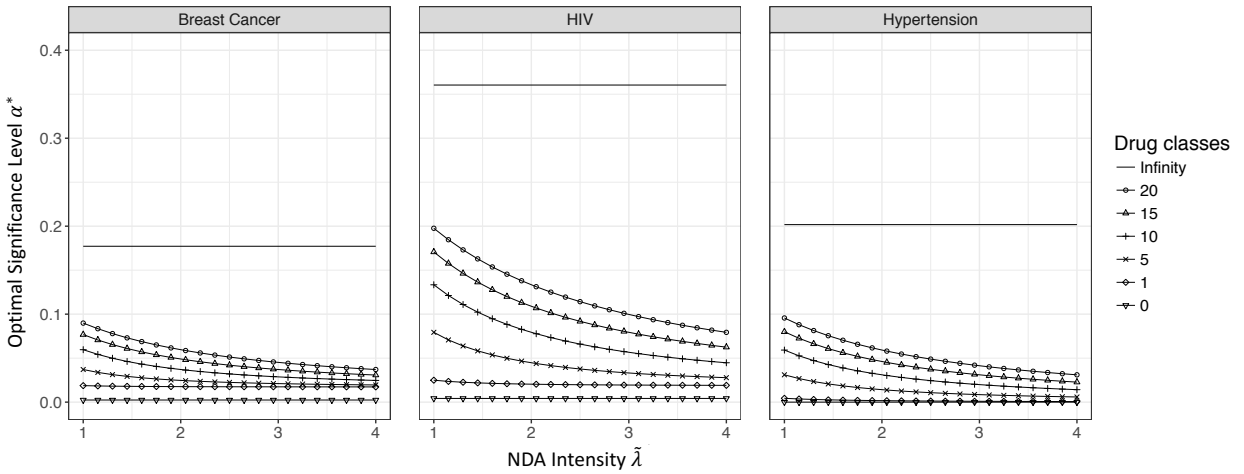
Sensitivity to NDA intensity. As suggested by Proposition 1, Figure 6 shows that the optimal approval policy α^* is decreasing in the rate of NDA submission $\tilde{\lambda}$. As more candidate drugs become eligible for approval, approving drugs is increasingly risky because we model health benefits as having diminishing marginal returns (due to obsolescence), but health costs as having constant

marginal returns. As the rate of NDA submission increases, the monetary losses and health costs of ineffective drugs eventually exceed the monetary gains and health benefits of effective drugs, so the optimal policy is to approve fewer drugs to avoid these potential negative outcomes. Figure 6 shows that the degree of sensitivity of the optimal policy depends on the value of K.

For small values of K, there is less capacity for additional drugs on the market, and thus for any given NDA intensity $\tilde{\lambda}$, it is likely that the market will be crowded ($\mathbb{E}[N_E] \approx K$). Under this scenario, there is little benefit in approving more drugs, and thus the NDA intensity has a minor effect on the optimal policy. In contrast, for large values of K, the optimal policy is more sensitive to the NDA intensity. When K is large, the market can support many effective drugs of different classes (e.g. ACE inhibitors, beta blockers, etc.) meaning that when the NDA intensity is low and the market is not crowded ($\mathbb{E}[N_E] << K$), the optimal policy is high and approves many drugs to fill the market, while when the NDA intensity is high and the market is crowded, the optimal policy is more conservative to avoid costs from ineffective drugs potentially gaining FDA approval.

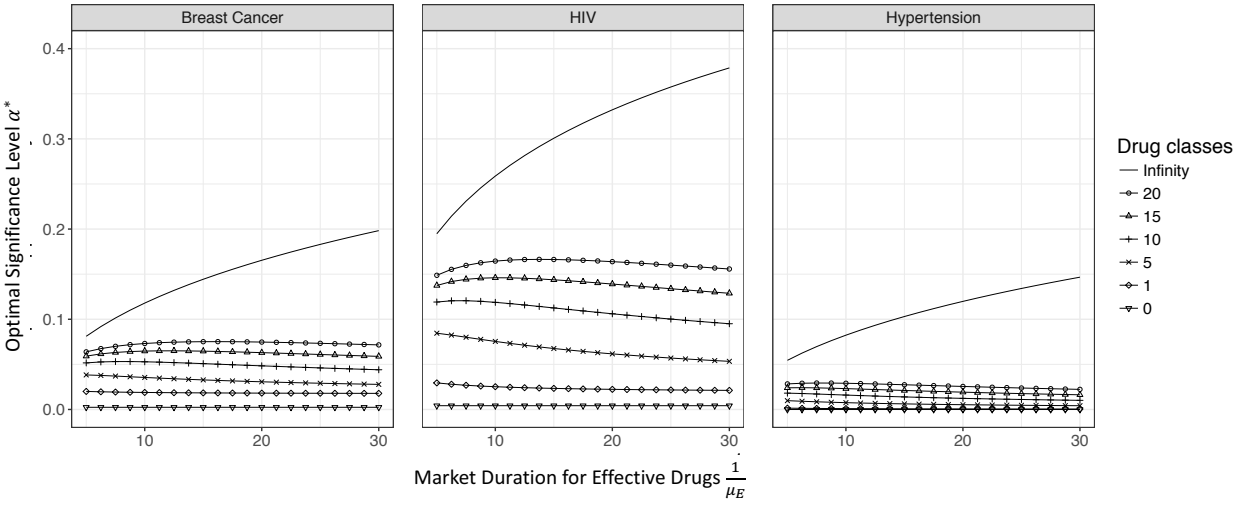
In the case that $K = \infty$ (the market for effective drugs is modeled as an $M/M/\infty$ queue), we see that the optimal policy is insensitive to the NDA intensity $\tilde{\lambda}$. This occurs because, when the post-approval phases are modeled as $M/M/\infty$ queues, the NDA intensity has the same marginal effect on health impacts and monetary values. Modeling the market for effective drugs in this manner has several drawbacks. First, the fact that the optimal policy is independent of the NDA intensity $\tilde{\lambda}$ means that this policy ignores several key characteristics of the drug development process (rate of clinical trial initiation, rate of clinical trial completion, and rate of attrition in the development process). Furthermore, the resulting approval policies are unrealistically high, with the model suggesting policies of $\alpha = 0.18$ (breast cancer), $\alpha = 0.36$ (HIV), and $\alpha = 0.20$ (hypertension).

Figure 6 Sensitivity of the optimal approval policy to the NDA intensity $\tilde{\lambda}$



Sensitivity to market duration. Recall from Proposition 1 that the optimal policy is non-monotonic with respect to the time spent on the market by effective drugs $1/\mu_E$. This behavior, more prominently seen for larger values of K in Figure 7, in which the curves first increase and then decrease, can be explained by the crowding of effective drugs on the market. Recall that high rates of market crowding result in more conservative approval policies because of diminishing marginal health benefits. When the time spent on the market $1/\mu_E$ is short, the expected number of effective drugs $\mathbb{E}[N_E(\alpha)]$ is small relative to the market capacity K. As $1/\mu_E$ increases, the market is still well below capacity and the monetary gains and health benefits of approving additional drugs supersede the monetary losses and health costs, so the optimal policy is to make it easier to approve drugs. However, as $1/\mu_E$ continues to increase, the market becomes saturated to the point where the monetary gains and health benefits of approving drugs no longer outweigh the monetary losses and health costs, and the optimal policy is to approve fewer drugs.

Figure 7 Sensitivity of the optimal approval policy to the market duration of effective drugs.



5.3. FDA Expedited Programs for Serious Conditions

Our model can also be used to examine the FDA's expedited programs for serious conditions: Accelerated Approval, Breakthrough Therapy, Fast Track, and Priority Review. These programs, whose qualifying criteria and features are summarized in Table 4, aim to benefit patients suffering from serious conditions by reducing the time to bring new drugs to market. The Accelerated Approval program allows drugs to be approved based on surrogate endpoints (e.g, tumor size, blood pressure), which can substantially reduce the time drug candidates spend in clinical trials, while the Priority Review program reduces the duration of the NDA review process from 10 months to 6 months. The Breakthrough Therapy and Fast Track programs are designed to expedite both

Table 4 Overview of FDA expedited programs.

Program	Qualifying Criteria	Features
Accelerated Approval (1992)	A drug that treats a serious condition and provides a meaningful advantage over available therapies and demonstrates an effect on a surrogate endpoint likely to predict clinical benefit.	Approval based on an effect on a surrogate endpoint.
Breakthrough Therapy (2012)	A drug that treats a serious condition and that preliminary evidence indicates may demonstrate substantial improvement on a clinically significant endpoint(s) over available therapies.	Intensive guidance on drug development; Rolling review.
Fast Track (1997)	A drug that treats a serious condition and nonclinical or clinical data demonstrate the potential to address unmet medical need.	Actions to expedite development/review.
Priority Review (1992)	A drug that treats a serious condition and, if approved, would provide a significant improvement in safety or effectiveness.	6-month FDA review (10-month standard)

Notes: Accelerated Approval was established under the 1992 Code of Federal Regulations, Breakthrough Therapy was established under the Food and Drug Administration Safety and Innovation Act of 2012, Fast Track was established under the Food and Drug Administration Modernization Act of 1997, and Priority Review was established under the Prescription Drug User Fee Act of 1992. Source: (FDA 2014a).

the clinical trial and review stages by allowing for frequent meetings between the FDA and drug developers, and by allowing for rolling review, in which portions of an NDA application can be submitted to the FDA at any time.

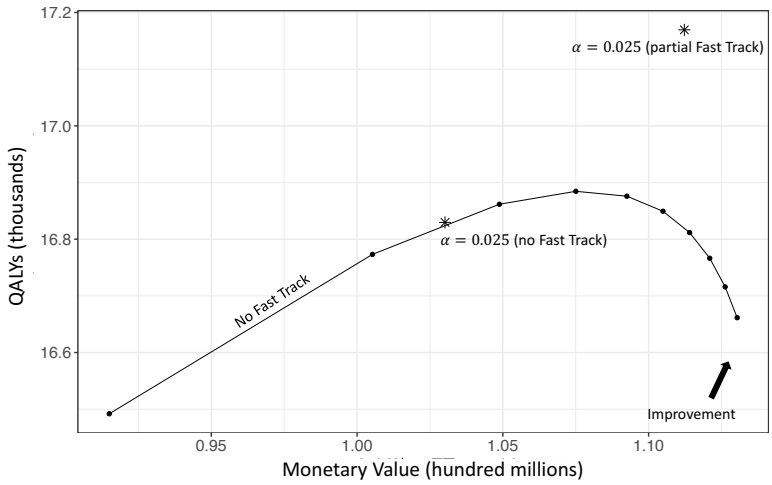
We focus our analysis on one expedited program (Fast Track), applied to one disease (breast cancer). The Fast Track program was chosen because of its impact on both the clinical trial and review durations, and because the Breakthrough Therapy program (which also affects both clinical trial and review duration) was only recently introduced in 2012. Breast cancer was selected because 48% of breast cancer drugs utilize the Fast Track program, compared to 35% of HIV drugs and only 1% of hypertension drugs (Kesselheim et al. 2015). We perform a counterfactual analysis by estimating the parameters of the FDA review process in the absence of Fast Track, and comparing the monetary value and QALYs obtainable under this scenario to the current system with Fast Track.

The Fast Track program is designed to reduce the time spent in clinical trials and NDA review, but not to affect any other aspect of the drug development and approval process (FDA 2014a). In our framework, this can be modeled as an increase in the clinical trial completion rate μ_{CT}. We assume that this is the only parameter that is affected by the presence of Fast Track, and that the monetary gains and losses, health benefits and costs, market durations, and effectiveness probability are unchanged. Although Fast Track may seem like an obvious improvement, its potential downsides include approving more ineffective drugs, lengthening the post-approval time with exclusive patent rights, and increasing the rate of drug obsolescence post-approval.

Let μ_0 denote the clinical trial completion rate under a system in which no drugs participate in the Fast Track program. Let μ_1 denote the clinical trial rate under a system in which all drugs participate in Fast Track, and let μ_{CT} denote the clinical trial rate under the current system, where

48% of breast cancer drugs use Fast Track and 52% do not. We denote the current system as having partial Fast Track. We use the value $\mu_{CT} = 0.08$, which is the clinical trial completion rate for breast cancer estimated in Section 5.2. We assume that the current duration of clinical trials, $\frac{1}{\mu_{CT}}$ is a weighted average of the duration of clinical trials under a system where all drugs use Fast Track and a system where none use this program, where the weights correspond to the proportion of breast cancer drugs utilizing Fast Track (48%) and not (52%). We also set $\frac{1}{\mu_1} = \frac{1}{\mu_{CT}} \cdot 0.95$ in accordance with a 2008 report by the Tufts Center for the Study of Drug Development that found that Fast Track reduced the total average clinical trial and review time by 5% with respect to all drugs.

Figure 8 Comparison of the monetary value ($C_{AE}\lambda_{AE}(\alpha) - C_{AI}\lambda_{AI}(\alpha) - C_{RE}\lambda_{RE}(\alpha)$) and QALYs achieved ($Q_E\mathbb{E}[N_E(\alpha)] - Q_I\mathbb{E}[N_I(\alpha)]$) under the current system (with partial fast track) and a system with no Fast Track.



Note. The approval policy α for the standard approval system varies from $\alpha = 0.01$ (the furthest left point) to $\alpha = 0.10$ (the furthest right point).

The curve in Figure 8 shows the trade-off in terms of monetary value and QALYs of varying α between 0.01 and 0.10 with no Fast Track. We also indicate the monetary value and QALYs achieved under the current approval system (with partial Fast Track) with a fixed policy of $\alpha = 0.025$. Compared to no Fast Track, the current system results in more monetary value and more QALYs when $\alpha = 0.025$. In other words, given a fixed approval policy, the addition of the Fast Track program dominates the approval process without this program. We observe that in the absence of Fast Track, no approval policy can achieve the QALYs obtainable under Fast Track. Eliminating Fast Track and increasing the approval threshold to $\alpha = 0.065$ generates similar monetary value to the current system (because a similar number of drugs are approved/rejected each year) but

significantly fewer QALYs because drugs spend more time in clinical trial and thus less time on the market.

We assume that Fast Track affects only the clinical trial completion rate μ_{CT}, this program could also result in a lower prior probability p of drug effectiveness. Shorter clinical trials mean less time to investigate interactions with other medicines or recruit different patient populations, while shorter FDA review times might mean less time to accurately evaluate clinical trial results. Assuming a fixed $\alpha = 0.025$, we find that for small changes in p, the current system continues to dominate the approval process with no Fast Track, both in terms of monetary value and QALYs. However, if p decreases to 0.86 (from $p = 0.912$) the current system no longer dominates in terms of monetary value, and if p decreases to 0.84 then an approval system with no Fast Track is strictly preferred.

6. Discussion

Our proposed queueing framework offers several insights into the FDA drug approval process, demonstrating how the pre-review process and post-approval market could influence a disease-specific FDA approval policy. Our model accounts for three key contributors to the shortfall of therapies available to treat some 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 help evaluate the impact of these programs on health benefits/costs and monetary gains/losses and, in the case of drugs that qualify for multiple programs, identify which program (or combination of programs) offers the largest societal benefit.

Disease-specific drug approval policies offer a fundamentally different way of addressing imbalances in the number of treatments available to patients. 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 is to ease approval standards for diseases with few drugs in the early stages of development (i.e., a low clinical trial arrival rate). By adjusting approval standards based on disease-specific characteristics, this approach has the potential to encourage pharmaceutical companies to reduce investment in certain diseases and focus development efforts on drugs that are likely to gain approval.

Our work is related to Montazerhodjat et al. (2017), who use Bayesian Decision Analysis to find the optimal statistical significance levels for oncology drugs, and compute an optimal level of 17.6% for breast cancer drugs— seven times higher than a traditional level of 2.5%. In comparison,

our model recommends a value of 4.6% for breast cancer. One driver of the discrepancy in these findings is the difference in how the post-approval market is modeled. We model ineffective drugs using an $M/M/\infty$ queue and, in an attempt to incorporate obsolescence, we model effective drugs using a collection of K $M/M/1/1$ queues, which results in diminishing marginal health benefits. As a result, our model places more weight on the costs of drug approval and thus recommends stricter approval standards. While our work accounts for obsolescence among drugs on the market, Montazerhodjat et al. (2017) ignores these effects and models effective and ineffective drugs in the same manner, which results in their comparatively less stringent approval policies.

While we focus our analysis on drug approval in the United States, our framework can be modified to model drug approval in other regions, such as Europe. Drugs developed in the United States and Europe both undergo clinical trial testing, but the review and approval processes differ substantially. In the United States, all drugs undergo centralized review by the FDA. In Europe, however, there are four possible paths to drug approval: a centralized process overseen by the European Medicines Agency, application to the regulatory body of a single European Union (EU) state, application for approval in all EU states following approval in one state, and independent application in multiple EU states (Van Norman 2016). A queueing framework such as the one presented in this work could be used to analyze the benefits of each approval pathway and to compare the European and American systems.

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 hypothesis testing 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 to ten 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. We aim to capture obsolescence and substitution by limiting the number of unique drug classes in the market, 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.

6.2. Future Work

Our study motivates several directions for future work. Currently, we model drug effectiveness as a binary variable, where drugs are either effective or ineffective, and we model drugs as having the same health impacts and monetary values in expectation. One extension is to model effectiveness as a continuous (or random) variable and/or model the health impacts and monetary values as random variables in order to account for heterogeneous responses of patients to a given treatment.

Another 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.

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 effectiveness of drugs while spurring development of novel therapeutics 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.

Ahmed P, Gardella J, Nanda S (2002) Wealth Effect of Drug Withdrawals on Firms and Their Competitors. *Financial Management* 31(3):21–41.

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.

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

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

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.

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) HIV/AIDS: HIV Prevalence Estimate. URL `https://www.cdc.gov/hiv/statistics/overview/index.html`, Accessed Jan 2018.

Chambers JD, Thorat T, Wilkinson CL, Neumann PJ (2017) Drugs Cleared Through The FDAs Expedited Review Offer Greater Gains Than Drugs Approved By Conventional Process. *Health Affairs* 36(8):14081415.

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.

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.

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

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.

Drummond MF, O'Brien B, Stoddart GL, Torrance GW (2003) *Methods for the Economic Evaluation of Health Care Programmes* (Oxford Medical Publications), 2nd edition.

Food and Drug Administration (2014a) Guidance for Industry Expedited Programs for Serious Conditions Drugs and Biologics . URL `https://www.fda.gov/downloads/Drugs/Guidances/UCM358301.pdf`, Accessed Feb 2018.

Food and Drug Administration (2014b) 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.

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.

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

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

Food and Drug Administration (2017a) Additions/Deletions for Prescription and OTC Drug Product Lists. URL `https://www.fda.gov/Drugs/InformationOnDrugs/ucm086229.htm`, Accessed Jan 2018.

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

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

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.

Food and Drug Administration (2017e) Multiple Endpoints in Clinical Trials Guidance for Industry . URL `https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm536750.pdf`, Accessed Feb 2018.

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

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.

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.

Kesselheim AS, Wang B, Franklin JM, Darrow JJ (2015) Trends in utilization of FDA expedited drug development and approval programs, 1987-2014: cohort study . *British Medical Journal* .

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.

Montazerhodjat V, Chaudhuri SE, Sargent DJ, Lo AW (2017) Use of Bayesian Decision Analysis to Minimize Harm in Patient-Centered Randomized Clinical Trials in Oncology. *JAMA Oncology* 3(9).

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

National Cancer Institute (2017) Cancer Stat Facts: Female Breast Cancer. URL `https://seer.cancer.gov/statfacts/html/breast.html`, Accessed Jan 2018.

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

Neumann PJ, Cohen JT, Weinstein MC (2014) Updating Cost-Effectiveness The Curious Resilience of the $50,000-per-QALY Threshold . *New England Journal of Medicine* .

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.

Sarkar SK, de Jong PJ (2006) Market Response to FDA Announcements. *The Quarterly Review of Economics and Finance* 46(4):586597.

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

Thomas DW, Burns J, Audette J, Carroll A, Dow-Hygelund C, Hay M (2016) Clinical Development Success Rates 2006-2015. Technical report, Biotechnology Innovation Organization, Biomedtracker, Amplion.

Tufts Center for the Study of Drug Development (2008) Fast track designations more than doubled during the last five years. Technical Report 5.

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

Van Norman GA (2016) Drugs and Devices: Comparison of European and U.S. Approval Processes. *JACC: Basic to Translational Science* 1(5):399–412.

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

Throughout this section, we suppress the dependence of various terms on α for readability and only explicitly note it when needed for clarity. For all derivatives, the variable of differentiation is α unless otherwise specified.

Proof of Theorem 1:

To show that $V(\alpha)$ is concave in α, we argue that $Q_E\mathbb{E}[N_E(\alpha)]$, $-Q_I\psi_I(\alpha)$, $C_{AE}\lambda_{AE}(\alpha)$, $-C_{AI}\lambda_{AI}(\alpha)$, and $-C_{RE}\lambda_{RE}(\alpha)$ are all concave functions of α, and thus the sum of concave functions is concave. Direct computation shows that $\mathbb{E}[N_E(\alpha)]$ is concave increasing in $\psi_E(\alpha)$ and that $\psi_E(\alpha)$ is concave in α. Thus $\mathbb{E}[N_E(\alpha)]$ is concave. Establishing concavity of the remaining terms is similarly straightforward. We note that in the case that $\alpha > 0$, $-C_{AI}\psi_{AI}(\alpha)$ and $-C_{RE}\lambda_{RE}(\alpha)$ are strictly concave in α and thus so is $V(\alpha)$. ■

Proof of Proposition 1:

By the Implicit Function Theorem, we have that

$$\frac{\partial \alpha^*}{\partial x} = -\frac{\frac{\partial V'(\alpha^*)}{\partial x}}{\frac{\partial V'(\alpha^*)}{\partial \alpha}} \tag{A.8}$$

where x is the parameter of interest. The fact that $V(\alpha)$ 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 $\frac{\partial V'(\alpha^*)}{\partial x}$. We use the equation

$$\begin{aligned} V'(\alpha) = &\left(Q_E \frac{\partial \mathbb{E}[N_E(\alpha)]}{\partial \psi_E}\frac{\partial \psi_E(\alpha)}{\partial \alpha} - Q_I \frac{\partial \mathbb{E}[N_I(\alpha)]}{\partial \psi_I}\frac{\partial \psi_I(\alpha)}{\partial \alpha}\right) WTP \\ &+ \left(C_{AE}\frac{\partial \lambda_{AE}(\alpha)}{\partial \alpha} - C_{AI}\frac{\partial \lambda_{AI}(\alpha)}{\partial \alpha} - C_{RE}\frac{\partial \lambda_{RE}(\alpha)}{\partial \alpha}\right) \end{aligned} \tag{A.9}$$

to find the sign of the effect of each parameter on α^*:

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial Q_E}\right) = \operatorname{sgn}\left(\frac{\partial \mathbb{E}[N_E(\alpha^*)]}{\partial \psi_E}\frac{\partial \psi_E(\alpha^*)}{\partial \alpha}\right) \geq 0 \tag{A.10}$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial Q_I}\right) = \operatorname{sgn}\left(-\frac{\partial \mathbb{E}[N_I(\alpha^*)]}{\partial \psi_I}\frac{\partial \psi_I(\alpha^*)}{\partial \alpha}\right) \leq 0 \tag{A.11}$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial C_{AE}}\right) = \operatorname{sgn}\left(\frac{\partial \lambda_{AE}(\alpha^*)}{\partial \alpha}\right) \geq 0 \tag{A.12}$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial C_{AI}}\right) = \operatorname{sgn}\left(-\frac{\partial \lambda_{AI}(\alpha^*)}{\partial \alpha}\right) \leq 0 \tag{A.13}$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial C_{RE}}\right) = \operatorname{sgn}\left(-\frac{\partial \lambda_{RE}(\alpha^*)}{\partial \alpha}\right) \geq 0 \tag{A.14}$$

$$\bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial \mu_I}\right) = \operatorname{sgn}\left(-Q_I\frac{\partial^2 \psi_I(\alpha^*)}{\partial \alpha \partial \mu_I}\right) \geq 0 \tag{A.15}$$

$$\begin{aligned} \bullet \operatorname{sgn}\left(\frac{\partial \alpha^*}{\partial \tilde{\lambda}}\right) = \operatorname{sgn}\Bigg(&\mathrm{WTP}\cdot Q_E\left(\frac{\partial \mathbb{E}[N_E(\alpha^*)]}{\partial \psi_E}\frac{\partial^2 \psi_E(\alpha^*)}{\partial \alpha \partial \tilde{\lambda}} + \frac{\partial^2 \mathbb{E}[N_E(\alpha^*)]}{\partial \psi_E^2}\frac{\partial \psi_E(\alpha^*)}{\partial \tilde{\lambda}}\frac{\partial \psi_E(\alpha^*)}{\partial \alpha}\right) \\ &-\mathrm{WTP}\cdot Q_I\frac{\partial^2 \psi_I(\alpha^*)}{\partial \alpha \partial \tilde{\lambda}} + C_{AE}\frac{\partial^2 \lambda_{AE}(\alpha^*)}{\partial \alpha \partial \tilde{\lambda}} - C_{AI}\frac{\partial^2 \lambda_{AI}(\alpha^*)}{\partial \alpha \partial \tilde{\lambda}} - C_{RE}\frac{\partial^2 \lambda_{RE}(\alpha^*)}{\partial \alpha \partial \tilde{\lambda}}\Bigg) \end{aligned} \tag{A.16}$$

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

$$\begin{aligned}\operatorname{sgn}\left(\tilde{\lambda}\frac{\partial\alpha^*}{\partial\tilde{\lambda}}\right) &= \operatorname{sgn}\left(\mathrm{WTP}\cdot Q_E\frac{\partial\mathbb{E}[N_E(\alpha^*)]}{\partial\psi_E}\frac{\partial\psi_E(\alpha^*)}{\partial\alpha} - \mathrm{WTP}\cdot Q_I\frac{\partial\psi_I(\alpha^*)}{\partial\alpha} + C_{AE}\frac{\partial\lambda_{AE}(\alpha^*)}{\partial\alpha}\right. \quad \text{(A.17)}\\ &\qquad \left. -C_{AI}\frac{\partial\lambda_{AI}(\alpha^*)}{\partial\alpha} - C_{RE}\frac{\partial\lambda_{RE}(\alpha^*)}{\partial\alpha} + \mathrm{WTP}\cdot Q_E\frac{\partial^2\mathbb{E}[N_E(\alpha^*)]}{\partial\psi_E^2}\psi_E(\alpha^*)\frac{\partial\psi_E(\alpha^*)}{\partial\alpha}\right)\\ &= \operatorname{sgn}\left(\mathrm{WTP}\cdot Q_E\frac{\partial^2\mathbb{E}[N_E(\alpha^*)]}{\partial\psi_E^2}\psi_E(\alpha^*)\frac{\partial\psi_E(\alpha^*)}{\partial\alpha}\right) \le 0 \quad \text{(A.18)}\end{aligned}$$

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

We claim that $\frac{\partial\alpha^*}{\partial\mu_E}$ and $\frac{\partial\alpha^*}{\partial p}$ are non-monotonic and that $\psi_E(\alpha^*) < 1$ is a sufficient condition to ensure that $\frac{\partial\alpha^*}{\partial\mu_E} \le 0$ and $\frac{\partial\alpha^*}{\partial p} \ge 0$. The proof of this is given by straightforward calculation of the sign of the desired derivatives:

$$\bullet\ \operatorname{sgn}\left(\frac{\partial\alpha^*}{\partial\mu_E}\right) = \operatorname{sgn}\left(-\frac{\tilde{\lambda}}{\mu_E^2}pe^{\Phi^{-1}(1-\alpha^*)\delta\sqrt{I_n}-\frac{\delta^2 I_n}{2}}\left(\frac{1-\psi_E(\alpha^*)}{(1+\psi_E(\alpha^*))^3}\right)\right) \quad \text{(A.19)}$$

$$\begin{aligned}\bullet\ \operatorname{sgn}\left(\frac{\partial\alpha^*}{\partial p}\right) = \operatorname{sgn}\Bigg(&\tilde{\lambda}e^{\Phi^{-1}(1-\alpha^*)\delta\sqrt{I_n}-\frac{\delta^2 I_n}{2}}\left(\mathrm{WTP}\cdot Q_E\left(1-\psi_E(\alpha^*)\right)+C_{AE}+C_{RE}\right) \quad \text{(A.20)}\\ &+\mathrm{WTP}\cdot Q_I\frac{\tilde{\lambda}}{\mu_I}+C_{AI}\tilde{\lambda}\Bigg)\end{aligned}$$

The condition $\psi_E(\alpha^*) < 1$ is sufficient to guarantee that $\frac{\partial\alpha^*}{\partial\mu_E} \le 0$ and $\frac{\partial\alpha^*}{\partial p} \ge 0$. ■

Proof of Proposition 2:

We begin by demonstrating that $\alpha_1^* \le \alpha_2^* \le \cdots \le \alpha_K^*$. To do this, we show that $V_K'(\alpha_{K+1}^*) \le 0$ for any $K \ge 1$. The concavity of $V_K(\alpha)$ will imply the desired inequality. Consider the following expression, where the notation $\mathbb{E}[N_E^K]$ and ψ_E^K is used to denote the expected number of effective drugs when there are K drug classes and the traffic intensity for each class, respectively:

$$\begin{aligned}V_K'(\alpha_{K+1}^*) - V_{K+1}'(\alpha_{K+1}^*) &= WTP\cdot Q_E\left(\frac{\partial\mathbb{E}[N_E^K(\alpha_{K+1}^*)]}{\partial\psi_E^K}\frac{\partial\psi_E^K}{\partial\alpha} - \frac{\partial\mathbb{E}[N_E^{K+1}(\alpha_{K+1}^*)]}{\partial\psi_E^{K+1}}\frac{\partial\psi_E^{K+1}}{\partial\alpha}\right) \quad \text{(A.21)}\\ &= -\frac{Q_E}{\mu_E}\frac{\partial\lambda_{AE}}{\partial\alpha}\frac{WTP}{(1+\psi_E^K)^2(1+\psi_E^{K+1})^2}\left(\frac{2\psi_E^K}{K+1}+\frac{(\psi_E^K)^2(2K+1)}{(K+1)^2}\right) \quad \text{(A.22)}\end{aligned}$$

From the optimality of α_{K+1}^*, we know that $V_{K+1}'(\alpha_{K+1}^*) = 0$, and thus noting that (A.22) is negative gives $V_K'(\alpha_{K+1}^*) \le 0$. As this holds for any K, we obtain the desired result.

Consider a system in which there is no market for effective drugs (i.e. $K = 0$). Applying the same argument as above gives

$$V_0'(\alpha_1^*) - V_1'(\alpha_1^*) = -WTP\cdot\frac{Q_E}{\mu_E}\frac{\partial\lambda_{AE}}{\partial\alpha}\frac{1}{(1+\psi_E^1)^2} \quad \text{(A.23)}$$

Noting that this expression is negative and that V_0' is concave in α, we see that

$$\alpha_0^* = 1-\Phi\left(\frac{1}{\delta\sqrt{I_n}}\log\left(\frac{1-p}{p}\frac{C_{AI}+WTP\cdot Q_I/\mu_I}{C_{RE}+C_{AE}}\right)+\frac{\delta\sqrt{I_n}}{2}\right) \le \alpha_1^* \quad \text{(A.24)}$$

where α_0^* is found by solving $V_0'(\alpha) = 0$.

Next, consider a system in which there are an infinite number of drug classes (i.e. $K = \infty$). We demonstrate that $\alpha_K^* \leq \alpha_\infty^*$. Note that $\mathbb{E}[N_E^K] = \frac{K\lambda_{AE}}{K\mu_E + \lambda_{AE}}$, and thus taking the limit of this expression as K goes to infinity gives $\mathbb{E}[N_E^\infty] = \frac{\lambda_{AE}}{\mu_E}$. Once again, we use the concavity of $V_K(\alpha)$ to establish the result. Consider the following expression:

$$V_K'(\alpha_\infty^*) - V_\infty'(\alpha_\infty^*) = WTP \cdot Q_E \left(\frac{\partial \mathbb{E}[N_E^K(\alpha_\infty^*)]}{\partial \psi_E^K} \frac{\partial \psi_E^K}{\partial \alpha} - \frac{\partial \mathbb{E}[N_E^\infty(\alpha_\infty^*)]}{\partial \alpha} \right) \tag{A.25}$$

$$= -WTP \cdot \frac{Q_E}{\mu_E} \frac{\lambda_{AE}}{\partial \alpha} \left(2\psi_E^K + (\psi_E^K)^2 \right) \tag{A.26}$$

By the optimality of α_∞^*, we have that $V_\infty'(\alpha_\infty^*) = 0$, and thus $V_K'(\alpha_\infty^*) \leq 0$. As a result, we have

$$\alpha_K^* \leq \alpha_\infty^* = 1 - \Phi \left(\frac{1}{\delta\sqrt{I_n}} \log \left(\frac{1-p}{p} \frac{C_{AI} + WTP \cdot Q_I/\mu_I}{WTP \cdot Q_E/\mu_E + C_{RE} + C_{AE}} \right) + \frac{\delta\sqrt{I_n}}{2} \right) \tag{A.27}$$

where α_∞^* can be found by solving $V_\infty'(\alpha) = 0$. ■

Proof of Proposition 3:

We begin by demonstrating that $V_K(\alpha_K^*) \leq V_{K+1}(\alpha_{K+1}^*)$, which first involves showing $V_K(\alpha) \leq V_{K+1}(\alpha)$ for all α. The following calculation shows that this is the case:

$$V_K(\alpha) - V_{K+1}(\alpha) = WTP \cdot Q_E \left(\frac{K\lambda_{AE}}{K\mu_E + \lambda_{AE}} - \frac{(K+1)\lambda_{AE}}{(K+1)\mu_E + \lambda_{AE}} \right) \tag{A.28}$$

$$= \frac{-WTP \cdot Q_E \cdot \lambda_{AE}^2}{(K\mu_E + \lambda_{AE})((K+1)\mu_E + \lambda_{AE})} \tag{A.29}$$

The series of inequalities $V_K(\alpha_K^*) \leq V_{K+1}(\alpha_K^*) \leq V_{K+1}(\alpha_{K+1}^*)$ completes this demonstration.

Next, we show 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 as follows:

$$V_K(\alpha) - V_\infty(\alpha) = WTP \cdot Q_E \left(\frac{K\lambda_{AE}}{K\mu_E + \lambda_{AE}} - \frac{\lambda_{AE}}{\mu_E} \right) \tag{A.30}$$

$$= -\frac{WTP \cdot Q_E \cdot \lambda_{AE}^2}{\mu_E(K\mu_E + \lambda_{AE})} \tag{A.31}$$

The remainder of the proof follows from the series of inequalities $V_K(\alpha_K^*) \leq V_\infty(\alpha_K^*) \leq V_\infty(\alpha_\infty^*)$. Next, we show $V_{K+1}(\alpha) - V_K(\alpha) > V_{K+2}(\alpha) - V_{K+1}(\alpha)$ by direct computation:

$$V_{K+1}(\alpha) - V_K(\alpha) - (V_{K+2}(\alpha) - V_{K+1}(\alpha)) \tag{A.32}$$

$$= WTP \cdot Q_E \left[\left(\frac{(K+1)\lambda_{AE}}{(K+1)\mu_E + \lambda_{AE}} - \frac{K\lambda_{AE}}{K\mu_E + \lambda_{AE}} \right) - \left(\frac{(K+2)\lambda_{AE}}{(K+2)\mu_E + \lambda_{AE}} - \frac{(K+1)\lambda_{AE}}{(K+1)\mu_E + \lambda_{AE}} \right) \right]$$

$$= \frac{2 \cdot WTP \cdot Q_E \cdot \mu_E}{((K+1)\mu_E + \lambda_{AE})\,((K+2)\mu_E + \lambda_{AE})\,(K\mu_E + \lambda_{AE})} \tag{A.33}$$

■

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.

- **Rate of clinical trial completion.** Let D_i denote the mean duration of Phase i trials, where $i =$I,II,III. We estimate $1/\mu_{CT}$ as $D_{\text{I}} + D_{\text{II}} + D_{\text{III}}$.

- **Rate of abandonment.** Recall for a drug that the probability of completing all three phases of clinical trials is given by

$$\mathbb{P}(\text{complete clinical trials}) = \frac{\mu_{CT}}{\mu_{CT} + \mu_{AB}} \tag{B.34}$$

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. Our estimate of the probability of completing clinical trials is the average of `Completed Phase III` across all drug interventions. Given our estimates of μ_{CT} and $\mathbb{P}$(complete clinical trials), we use equation B.34 to solve for our estimate of μ_{AB}.

- **Rate of clinical trial initiation and NDA submission.** In order to estimate the NDA submission rate λ and clinical trial initiation rate $\tilde{\lambda}$, we first note that the rate $\lambda_{AE} + \lambda_{AI}$ at which drugs are approved is the product of the rate at which NDAs are submitted $\tilde{\lambda}$ and the probability that a submitted NDA is approved, $\mathbb{P}$(Approve NDA). We estimate the average rate $\lambda_{AE} + \lambda_{AI}$ at which drugs were historically approved using exhaustive lists of drugs approved to treat a disease (Tables B2 - B4), and we use estimates for $\mathbb{P}$(Approve NDA) from Thomas et al. (2016). Using our estimates of $\lambda_{AE} + \lambda_{AI}$ and $\mathbb{P}$(Approve NDA), we obtain our estimate of $\tilde{\lambda}$ as

$$\tilde{\lambda} = \frac{\lambda_{AE} + \lambda_{AI}}{\mathbb{P}(\text{Approve NDA})}. \tag{B.35}$$

The rate at which drugs begin clinical trials λ is then estimated as

$$\lambda = \frac{\tilde{\lambda}}{\mathbb{P}(\text{Complete clinical trials})}. \tag{B.36}$$

- **Clinical trial information.** The clinical trial information $\delta\sqrt{I_n}$ is estimated by assuming the statistical power of the trial—the probability of approving a drug conditional on the drug being effective (given by π_{AE}/p)—is 90%, given a traditional statistical significance level of $\alpha = 2.5\%$. Mathematically, our estimate $\delta\sqrt{I_n}$ is chosen to satisfy

$$.90 = 1 - \Phi\left(\Phi^{-1}(1-0.025) - \delta\sqrt{I_n}\right). \tag{B.37}$$

- **Effectiveness probability.** To estimate the prior probability p that a drug is effective, we select the value of p that makes the probability of approving a drug in our model equal to the estimated probability that an NDA is approved, assuming $\alpha = 2.5\%$. Thus our estimate p satisfies

$$\begin{aligned}\mathbb{P}(\text{Approve NDA}) &= \pi_{AE}(\alpha) + \pi_{AI}(\alpha) &\text{(B.38)}\\ &= \left[1 - \Phi\left(\Phi^{-1}(1-0.025) - \delta\sqrt{I_n}\right)\right] p + (1-0.025)p. &\text{(B.39)}\end{aligned}$$

Monetary Values. To estimate the monetary values C_{AE}, C_{AI}, and C_{RE}, we multiply the median pharmaceutical market capitalization $Market\ Cap$ by the percent change in market capitalization as a result of approving effective, approving ineffective, and rejecting effective drugs, respectively. We use published estimates from Sarkar and de Jong (2006) and Ahmed et al. (2002) of percent abnormal market returns at the time of initial review $r_{initial}$, the time a drug is announced as approvable $r_{approvable}$, the approval announcement day $r_{approval\ day}$ (or the rejection announcement day $r_{rejection}$), the day after the approval announcement $r_{day\ after\ approval}$, and following market withdrawal $r_{withdrawal}$. We combine these values with the median pharmaceutical market capitalization to obtain the following monetary value estimates:

$$C_{AE} = (r_{initial} + r_{approvable} + r_{approval\ day} + r_{day\ after\ approval}) \cdot Market\ Cap \tag{B.40}$$

$$C_{AI} = C_{AE} - (r_{withdrawal}) \cdot Market\ Cap \tag{B.41}$$

$$C_{RE} = (r_{initial} + r_{approvable} - r_{rejection}) \cdot Market\ Cap \cdot p. \tag{B.42}$$

Note that the probability p that a drug is effective appears in our estimate for C_{RE}, but not in our estimates for C_{AE} or C_{AI}. In the case of approved drugs, we assume that it is possible to distinguish the monetary value of effective and ineffective drugs using the market reaction to drug withdrawals. However, in the case of rejected drugs, this differentiation is not possible, so instead we multiply the change in pharmaceutical market capitalization by the probability that a drug is effective.

Table B1 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); Department of Health and Human Services (2016); World Health Organization (2016); Agency for Healthcare Research and Quality (Townsend et al. 2011).

Table B2 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

Sources: National Cancer Institute (2016), Food and Drug Administration (2017a)

Table B2 FDA-approved breast cancer drugs (continued).

Drug (Brand Name)	Approval	Drug Class
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	
Ribociclib (Kisqali)	Mar 2017	
Neratinib maleate (Nerlynx)	July 2017	
Abemaciclib (Verzenio)	Sep 2017	
Olaparib (Lynparza)	Jan 2018	
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), Food and Drug Administration (2017a)

Table B3 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	
Emtricitabine and Tenofovir Alafenamide (Descovy)	Apr 2017	
Dolutegravir & Rilpivirine (Juluca)	Nov 2017	
Bictegravir & Emtricitabine & Tenofovir & Alafenamide (Bictegravir)	Feb 2018	
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), Food and Drug Administration (2016, 2017a)

Table B4 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 B4 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: Food and Drug Administration (2017a)

Table B5 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.
Sources: Avastin - Drugsite Trust (2018a), Hivid - Food and Drug Administration (2018), International Association of Providers of Aids Care (2017), Ticrynafen - Manier (1982), Posicor - Bradbury (1998), Valturna - Drugsite Trust (2018b), Food and Drug Administration (2017b)

References

Ahmed P, Gardella J, Nanda S (2002) Wealth Effect of Drug Withdrawals on Firms and Their Competitors. *Financial Management* 31(3):21–41.

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

Bradbury J (1998) Posicor withdrawn voluntarily from market by Roche . *The Lancet* 9118(351):1791.

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.

Drugsite Trust (2018a) Avastin Approval History. URL `https://www.drugs.com/history/avastin.html`, Accessed Jan 2018.

Drugsite Trust (2018b) Valturna Approval History. URL `https://www.drugs.com/history/valturna.html`, Accessed Jan 2018.

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

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

Food and Drug Administration (2017b) FDA Drug Safety Communication: New Warning and Contraindication for blood pressure medicines containing aliskiren (Tekturna). URL `https://www.fda.gov/Drugs/DrugSafety/ucm300889.htm`, Accessed Jan 2018.

Food and Drug Administration (2018) HIV/AIDS Historical Time Line 1995-1999. URL `https://www.fda.gov/ForPatients/Illness/HIVAIDS/History/ucm151079.htm`, Accessed Jan 2018.

International Association of Providers of Aids Care (2017) Zalcitabine (Hivid, ddC). URL `http://www.aidsinfonet.org/fact_sheets/view/412`, Accessed Jan 2018.

Manier JW (1982) Hepatotoxicity associated with ticrynafen–a uricosuric diuretic. *The American journal of gastroenterology* 77(6).

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.

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

Sarkar SK, de Jong PJ (2006) Market Response to FDA Announcements. *The Quarterly Review of Economics and Finance* 46(4):586597.

Thomas DW, Burns J, Audette J, Carroll A, Dow-Hygelund C, Hay M (2016) Clinical Development Success Rates 2006-2015. Technical report, Biotechnology Innovation Organization, Biomedtracker, Amplion.

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.

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