On the Economic Infeasibility of Personalized Medicine, and a Solution

Marina Chiara Garassino
University of Chicago Comprehensive Cancer Center

Kunle Odunsi
University of Chicago Comprehensive Cancer Center

Marciano Siniscalchi
Northwestern University

Pietro Veronesi
University of Chicago, NBER, and CEPR

May 20, 2025

Abstract

Technological advances and genomic sequencing opened the road to personalized medicine: specialized therapies targeted to patients displaying specific molecular alterations. For instance, targeted therapies are now available for 50% of lung cancer patients—with some alterations affecting less than 1% of patients—greatly increasing life expectancy. In an investment model of drug development, we show that current institutions mandating experimentation and approval of individual therapies eventually disincentivize investments in personalized medicine as researchers identify increasingly rare alterations. Recent AI-based technologies, such as AlphaFold3, make personalized medicine viable when regulatory approval regards the *process* for drug discovery rather than individual therapies.

Keywords: targeted therapies, precision medicine, genetic alterations.

JEL codes: I11, I18, O32, O38

---

* Veronesi acknowledges financial support from the Fama-Miller Center for Research in Finance and by the Center for Research in Security Prices at the University of Chicago Booth School of Business. For useful comments, we thank workshop participants at the Chicago Fed, Bocconi University, University of Chicago, the 2025 Health Economics Conference at the Becker Friedman Institute, and the 2025 WashU Conference on the Economics and Finance of Healthcare and Medicine. We thank Eric Budish, Alex Frankel, Lars Hansen, Andrew Lo, and Tong Liu (discussant) for insightful comments. We also thank Salem Soin-Voshel for excellent research assistance. Errors are our own. The views expressed in this paper are our own and do not necessarily reflect the views of our respective employers. Siniscalchi declares that he has no relevant or material financial interests that relate to the research described in this paper. Garassino, and Veronesi's partner, declares that she serves as paid consultant of several pharmaceutical companies, and also serves on their steering committees for clinical trials. Odunsi is supported by P30CA014599 grant from the NIH and is a co-founder of Tactiva Therapeutics. Contact information: Marina Garassino: mgarassino@bsd.uchicago.edu; Kunle Odunsi: odunsia@bsd.uchicago.edu; Marciano Siniscalchi: marciano@northwestern.edu; Pietro Veronesi: pietro.veronesi@chicagobooth.edu.

1. Introduction

The publication of the initial sequencing of the human genome (International Human Genome Sequencing Consortium, 2001) opened up the road to modern personalized medicine: that is, delivering the right drug to the right patient based on their genomic profile. Technological advances over the last twenty-five years have indeed enabled the development of effective treatments for diseases that affect increasingly smaller subgroups of patients. As an example, highly specialized targeted therapies are now available for about 50% of all lung-cancer patients, including patients who never smoked. Such therapies are significantly more effective than non-targeted alternatives: for instance, for certain types of lung cancer due to specific and rare genetic alterations such as ALK and ROS1, targeted therapies have extended life expectancy from a few months to several years. Promising advances in personalized medicine are also occurring in other areas of medicine, such as other types of cancer, hematological, cardio-vascular, immunological and neurodegenerative disorders, and rare genetic diseases. Such results warrant a degree of optimism that a new era of personalized medicine will bring about significant improvements for patients affected by devastating illnesses. However, the highly targeted nature of personalized medicine introduces new challenges. The very notion of personalized medicine implies that, theoretically, each patient can become a unique case. By way of contrast, the current regulatory infrastructure is founded on statistical testing for safety and effectiveness, and thus assumes the availability of sufficiently large subject pools.

This paper uses a simple model of pharma companies' investment decisions to study the economic sustainability and the incentives to personalized drug discovery as scientific advances identify ever more specific molecular alterations that, on one hand, can be treated with greater effectiveness, but, on the other hand, have a necessarily lower incidence in the population. Specifically, we focus on the impact of the current regulatory infrastructure on firms' decisions. Regulatory agencies such as the Federal Drug Administration (FDA), the European Medicine Agency (EMA), and so on, require that new drugs be validated as more effective than previously available treatments through randomized control trials (RCT).[1] This process involves different phases (see Section 3.), but the basic structure is as follows. Patients are randomly assigned to a "treatment" group, which receives the new proposed (targeted) therapy, or a "control" group, which receives the best currently available therapy. Patients are then followed for years, and survival statistics are used to compare "success rates" in the two groups. The new treatment is approved if the difference in success rates is medically as well as statistically significant.

[1] Over the years, the regulatory agencies have relaxed some of the strict requirements and allowed for accelerated approvals for some specific cases. See https://www.fda.gov/drugs/development-approval-process-drugs for details about the FDA approval process and designations.

We show that, under reasonable regularity conditions, for a given disease (e.g. lung cancer), as the number of treatable molecular alterations increases, with each alteration affecting an increasingly smaller fraction of patients, the probability of a successful RCT decreases. The reason is that, even for targeted treatments that are meaningfully superior to the current standard of care, it becomes increasingly difficult to secure large enough samples to meet the required significance threshold. Moreover, holding the prices of targeted therapies fixed, the expected profits for the pharma industry as a whole decrease to zero when the number of treatable molecular alterations increases, due to lower expected profits and higher overall costs of running RCTs. Since currently pharma companies are primarily responsible for the discovery of new targeted therapies, these results imply that eventually there will be no incentive to develop targeted therapies for newly discovered biomarkers. In principle, this can be avoided by increasing the prices of targeted therapies as the number of treatable molecular alterations increases. Indeed, break-even prices increase as the number of alterations increases, and are unbounded in the limit, regardless of parameter values. This eventually makes the overall costs of drug discovery unsustainable for society. In other words, the current regulatory regime makes personalized medicine infeasible in practice.

To assess magnitudes, we calibrate the model to data from lung cancer research and patient care. Lung cancer is a good example of the progress made in personalized medicine, as it has gone through several breakthroughs in the last 20 years (see the following section). In particular, researchers have discovered that many types of lung cancer are due to genetic alterations that affect as low as 0.1% or even 0.05% of lung cancer patients. In our calibration, we find that the pharma industry stands to make high profits when the number of alternations is limited; this explains the explosion in research and R&D investments in targeted therapies over the last two decades. However, the present value of such profits rapidly declines to zero as the number of alterations increases. Indeed, even for R&D costs that are conservatively far lower than those reported by pharma companies, the present value of future profits reaches zero when the number of alterations is just over 60—equivalently, when alterations affect about 1.7% of the patient population. This is the case even considering tax credits and other incentives, such as those provided by the Orphan Drug Act in the U.S. We also calculate the societal costs of lung cancer treatment at break-even prices, and find that, even for conservatively low costs of R&D, they run in the trillions of dollars per year as the number of alterations increases, affecting a progressively smaller—yet realistic—number of patients.

Our analysis indicates that pharma companies will have little to no incentive to innovate and find new drugs for alterations that affect an ever decreasing size of the population, despite the fact that, in the aggregate, such alterations cover the entire population of lung-cancer patients. Indeed, we find empirical evidence that the cumulative number of clinical

trials that are run by pharma companies over a ten-year span since discovery of an alteration significantly decreases with the percentage of the population affected by that alteration.

Our model, however, also indicates that a regulatory framework in which regulatory agencies approve the *process* for drug discovery instead of individual drugs leads to a sustainable system that encourages drug development. Recent technological advances make such a process realistic. In brief, a "personalized drug discovery process' (PDDP) would comprise five steps (see section 6.1. for details): (A) the identification of driver alterations, which can now be done through many existing filters; (B) Rebuilding of the altered protein, which can now be accomplished by using machine learning methods and AI, such as Google DeepMind AlphaFold3; (C) compound matching, which involves using supercomputers to find, out of billions of possible compounds, one that can "block" the altered protein and hence inhibit the growth of the tumor; (D) Drug discovery, which uses the knowledge of molecular structure of existing safe drugs and machine learning methods to ensure that the compound identified in previous steps can be absorbed and tolerated by humans; (E) Drug production.

Very recent research in biochemistry and bioinformatics shows that steps (A) – (D) can in fact be achieved effectively in silico (i.e., through simulation) very rapidly, even within one month (see e.g.Ren et al., 2023). A personalized treatment would start with a patient being tested to detect the presence of alterations (step A), and continue with steps (B) – (D). If a new personalized drug is found, then it is produced — step (E). Otherwise, the patient is treated using current available therapies. Overall, PDDP involves spending about a month or less to find the proper personalized, safe drug.

Building on our calibration results, we show that PDDP is economically sustainable. Once approved, the costs of the process of drug discovery does not depend on the number of alterations, and there is no need to run expensive RCTs for every alteration. However, the incentive of pharma companies to switch to PDDP depends on the probability of finding a personalized drug through steps (A) – (D) above, and the survival probability delivered by the newly discovered drug. We show that pharma companies have incentives to adopt the PDDP over the previous existing therapies if the annual profit margins of PDDP therapies range between 70% (when annual survival rate is 60%) and 12% (when annual survival rate reaches 90%) of typical current annual profit margins of targeted therapies, which will bring substantial savings to lung cancer patients.

Our paper is related to several strands of literature, both in the medical field and in the economics field. In the medical field, our work is related to both the literature documenting the progress in targeted therapies (see e.g Hendriks et al., 2024 for a recent review), the costs of targeted therapies (see e.g. Leighl et al., 2021), and the recent work on the use of computers for protein projection and compound discovery (see e.g. Ren et al., 2023). However, none of

this previous work appears to have tackled the profitability of drug discovery as the number of alterations increase. To the best of our knowledge, the personalized drug discovery process we analyze is also novel.

In economics, a relatively recent literature examines the problem of regulating the approval of new products such as drugs: e.g. Carpenter and Ting (2007), Henry and Ottaviani (2019), and Henry, Loseto, et al. (2022). This literature emphasizes that the regulator faces a trade-off between inducing an appropriate level of experimentation, which is socially beneficial, and preventing harmful products from reaching consumers. More broadly, this approach is related to the analysis of strategic experimentation (e.g. Guo, 2016; McClellan, 2022) and (dynamic) persuasion (Kamenica et al., 2011; Honryo, 2018). Our paper is also related to the recent literature on investments of pharma companies in drug development. Frankel, Krieger, Li, and Papanikolaou (2024) compares novel or "breakthrough" drugs vs. incremental drugs, and finds that firms have stronger incentives to invest in the latter, despite the fact that the former yield greater knowledge spillovers. Budish, Roin, and Williams (2015) shows that commercialization lags and patent terms distort firms' incentives in R& D relative to the social optimum; for cancer drugs, this leads, for instance, to limited investment in preventative treatments, and overinvestment in drugs for late-stage cancers. One of their proposed solutions involves targeting surrogate (non-survival) outcomes in the approval process, which in our model is comparable to an increase in the probability of approval. Our paper differs in that it focuses on the details of the approval process, in particular as it applies to personalized medicine.

Our approach is complementary. We take a stylized form of the current regulatory regime as given, and study the investment problem faced by a pharma company, and by the pharma industry as a whole. One can interpret our main negative result as showing that the current regime cannot be part of an incentive-compatible mechanism that implements the first best. We show that even when the initial R&D effort has been undertaken and has produced an effective treatment, the firm is not willing to carry out the expensive testing process once the number of genetic alterations is large enough. This is true even when the firm knows (but cannot otherwise credibly communicate) that the success probability is strictly higher than the previous standard of care —- a case in which the first-best outcome is approving the new treatment. Thus, the current regime does not implement the first best. However, the alternative approval mechanism, the "personalized drug discovery process," sustains the approval of new treatments for a (much) wider range of success probabilities.

2. Targeted Therapies in Lung Cancer

According to GLOBOCAN 2022, lung cancer is the most commonly diagnosed cancer, accounting for 12.1% of all cancers, and the leading cause of cancer-related deaths, responsible for 18.8% of all cancer-related fatalities, with an estimated 1.8 million deaths worldwide.[2]

Lung cancer is also one of the most aggressive tumors, with a very high mortality rate. Until the early 2000s, the only available medical treatment was chemotherapy. Over the past fifteen years, however, major advances in understanding the biology of tumors have led to the advent of target therapies and immunotherapy. These discoveries brought about a major increase in patient survival expectancy. For instance, in metastatic patients, median survival went from months to years.

In 2004, for the first time, lung cancer researchers identified specific genomic mutations in a portion of a gene called EGFR in the tumor of the patients. Patients harboring these alterations in the tumor were very sensitive to a particular class of agents, called EGFR tyrosine inhibitors. This discovery opened up the era of personalized medicine in lung cancer, as well as in many other tumors. Through randomized clinical trials, in the following few years, medical research demonstrated that these drugs were superior to the "one-size-fits-all approach" of chemotherapy. Many other targets were found in lung cancer and in other tumors. Conversely, those patients without these alterations were totally resistant to the new drugs. This finding created a paradigm shift: for each altered gene there was a corresponding tyrosine inhibitor. Such alterations are typically more prevalent in the population of patients who never smoked, which is about 15% of the overall lung cancer population in U.S..

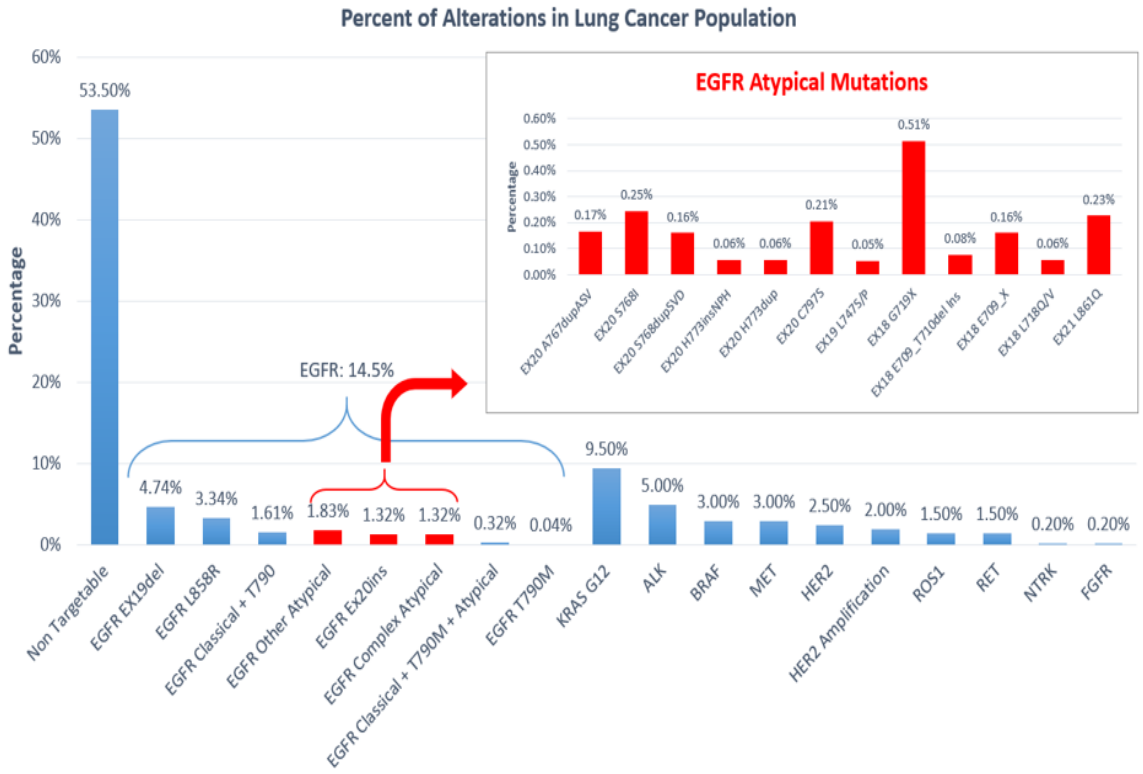
Following the discovery of EGFR mutations in 2004, several other targets were identified. ALK and ROS1 fusions were identified in 2007, followed by BRAF, RET, HER2, MET, NTRK, KRAS G12, EGFR Exon 20, and NRG1. Figure 1 indicates the frequency of alterations in lung cancer patients.[3] About 53.5% of patients with lung cancer do not have any specific alteration for which target treatment is available. Approximately 14.5% have the EGFR alteration, about 9.5% have the KRAS G12 mutation, and so on. A few of these alterations, such as ROS1 or RET, affect only 1% to 2% of the population; and others still, such as NTKR and FGFR, only affect less than 1% of the population.

Moreover, even within pre-defined groups of alterations, there is still some heterogeneity that can be responsible for a different prognosis and a different sensitivity to drugs. For

[2]See https://gco.iarc.fr/today/en/dataviz/pie?mode=cancer&group_populations=1&populations=900&age_end=17&types=1 accessed on 1/20/2024.

[3]These estimates are from Taha et al. (2021), who report ranges of frequencies. In particular, they report that between 39% and 68% of patients with lung cancer do not have any specific alteration; between 12% and 17% have a genetic alteration called EGFR, between 7% and 12% they have the alteration KRAS G12, and so on. The figure reports the midpoints of the respective ranges for simplicity.

Figure 1: **Percentage of Alterations in Lung Cancer Patients**

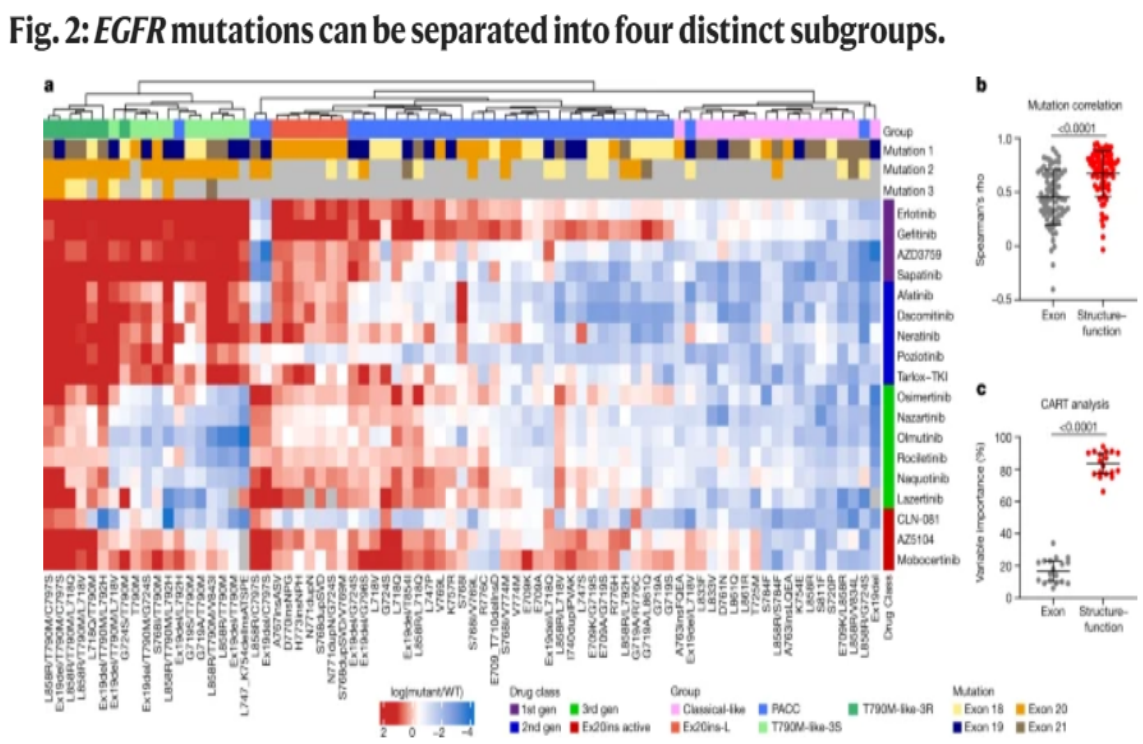


Source: Taha et al. (2021), Robichaux et al. (2021)

instance, the red panel in Figure 1 shows the percentages of specific atypical EGFR alterations, which have different degrees of sensitivity to EGFR tyrosine kinase inhibitors, even within the same EGFR family. The frequencies of some of these atypical alterations are as low as 0.05%. Figure 2, taken from Figure 2 of Robichaux et al. (2021), shows the vast number of possible alterations even within the EGFR group, and the relative efficacy (or lack thereof) of known therapies across alterations. Indeed, Roskoski Jr (2024) reports that, as of 2024, there are only 80 FDA-approved drugs that target about two dozens of different protein kinases, which still represent "a small fraction of the 518-member protein kinase enzime superfamily." While there are currently 180 clinical trials worldwide studying orally effective protein kinase inhibitors, there are still a vast number of potential alterations and alterations that are not researched on, or grouped in a pre-defined category.

Each of these alterations acts differently on the cellular system, and thus requires a specific targeted therapy that act on each of the specific alterations. The number of new targeted therapies has exploded in recent years, as each different treatment targets the same alteration with different specific compounds (see Figure 3). The development of targeted drugs has not been confined to the treatment of lung cancer: it has extended to many other tumors, such as melanoma, thyroid cancer, gastric cancer, etc. In addition, over time, these

Figure 2: **EGFR alterations**



(Source: Figure 2 in Robichaux et al. 2021)

Figure 3: **Drug Discovery Progression**



tumors have become resistant to the first generation of drugs. Consequently, new generations of drugs have been developed to overcome the resistance to the previous inhibitors. Table 1 shows the list of FDA–approved treatments for various alterations. This table indicates an acceleration in the number of available treatments over the last decade.[4] Numerous randomized control trials for new treatments are ongoing.

[4]The table reports the initial approval time, as some drugs first receive accelerated approval before final approval. Others gain approval for certain indications first, later expanding to additional uses, sometimes with other treatments.

Table 1: **FDA Approved Drugs for Targeted Therapies**

Agent	Control	Manufacturer	FDA Approval	Price per week	mPFS (months)	R&D Cost (millions)
Panel A: First Line						
EGFR (12% - 17%)						
Gefitinib	vs. chemo	AstraZeneca	2003*	$1,715	10.8 vs 5.4	
Erlotinib	vs. chemo	Roche	2004*	$1,785	9.7 vs. 5.2	$124
Afatinib	vs. chemo	Boehringer	2013	$2,733	12.4 vs. 6.8	
Dacomitinib	vs. Gefitinib	Pfizer	2018	$3,815	14.7 vs. 9.2	
Osimertinib	vs. Gefitinib	AstraZeneca	2018	$3,966	18.9 vs. 10.2	$594
ALK (3% - 7%)						
Crizotinib	vs. chemo	Pfizer	2011	$5,200	7.7 vs. 3.0	
Ceritinib	vs. chemo	Novartis	2017	$5,346	16.6 vs. 8.1	$1,095
Alectinib	vs. Crizotinib	Roche	2015	$4,299	34.8 vs. 10.9	$409
Brigatinib	vs. Crizotinib	Ariad	2020	$4,699	24.0 vs. 11.1	$491
Lorlatinib	vs. Crizotinib	Pfizer	2021	$4,953	NR vs. 9.1	
ROS1 (1% - 2%)						
Crizotinib	vs. chemo	Pfizer	2016	$5,200	18.4 vs. 8.6	
Entrectinib	SA	Roche	2019	$4,683	19.0	$451
Repotrectinib	SA	Brystol Myers Squibb	2023		35.7	
RET (1% - 2%)						
Selpercatinib	vs. chemo	Eli Lilly	2022	$4,557	24.8 vs. 11.2	$511
Pralsetinib	SA	Roche/Blueprint	2023	$5,208	NR	$260
BRAF (1% - 5%)						
Dabrafenib tramentinib	SA	Novartis	2017	$3,325	14.4	$1,239[1]
MET ex 14 (2% - 4%)						
Capmatinib	SA	Novartis	2020	$5,684	12.4	$561
Tepotinib	SA	Merck	2021	$6,315	11.0	
NTRK (<1%)						
Larotrectinib	SA	Bayer	2018	$9,114	24.6	$809
Entrectinib	SA	Roche	2019	$4,683	11.2	$451
EGFR Exon 20 (4.7%)						
Amivantamab - chemo	vs. chemo	J&J	2021*	$6,654	11.4 vs. 6.7	
Panel B: Second Line						
EGFR (12% - 17%)						
Osimertinib	vs. chemo	AstraZeneca	2018	$3,966	10.1 vs. 4.4	$594
Amivantamab/chemo	vs. chemo	J&J	2021	$6,654	8.2 vs. 4.2	
KRAS G12C (7% - 12%)						
Sotorasib	SA	Amgen	2021	$4,935	6.8	
Adagrasib	SA	Mirati	2022	$5,124	6.5	
HER2 (1% - 4%)						
Trastuzumab Deruxtecan	SA	Roche	2022	$4,104	8.2	

Source: National Cancer Institute, FDA, Drugs.com, Taha et al. (2021). Prices are from drug.com, accessed in Spring 2023. Dosage per week are from the respective clinical trials. R&D costs are from Henderson et al. (2023), supplemental table S3. Footnotes: * denotes first FDA approval. [1] R&D cost for Trametinib itself, approved in 2013.

EGFR is a useful representative example of the drug discovery process in the new world of targeted therapies, and its development over the years is similar to that for other targets.

The majority of the activating EGFR mutations are concentrated in four exons[5], 18, 19, 20 and 21. Those in exon 19 and exon 21 are called "typical or classical mutations." These mutations appear in nearly 70% of cases sensitive to the first, second, and third generation of tyrosine kinase inhibitors (TKIs). However, the mutations can also happen in other areas, such as exon 18 and 20: these are called "atypical mutations," and display heterogeneous response to EGFR TKIs. A typical example are the mutations (insertions) in exon 20, for which new drugs are currently under development. In 2004, after the discovery of EGFR mutations, it was found that these tumors were very sensitive to the drugs Gefitinib and Erlotinib, known as first generation EGFR TKIs. A few years later, the second-generation inhibitors were commercialized; these include Afatinib and Dacomitinib. A further advance was the discovery of mechanisms of resistance to the first- and the second-generation TKIs. This led to the development of Osimertinib, which was shown to be superior to Gefitinib and Erlotinib in randomized clinical trials. As a result, the earlier-generation TKIs ceased to be used in the majority of countries.[6] Most recently, research has focused on approaches to targeting resistance to Osimertinib, and many compounds are now under investigation. For atypical mutations, representing ultra-rare populations resistant to Osimertinib, research is still ongoing, and recent results have shown that the drug Amivantanab, combined with chemotherapy, can be superior to chemotherapy alone in Phase III trial. Similar drug development happened also for patients harboring a tumor with ALK fusion treated with the corresponding ALK inhibitors.

The developments discussed in the previous paragraph highlight the increasingly specificity of targeted therapies, even within a specific group of mutations, such as EGFR. Figure 1 shows the percentage of specific mutations within the EGFR mutation group. Recall that EGFR mutations affect about 14.5% of the lung cancer population. Of these, about 32.7% (4.7% of lung cancer population) have the EX19del mutation, another 23% (3.3% of lung cancer population) have the L858 mutation, and so on. Each targeted therapy is developed to attack different subgroups of EGFR mutations, making the percentage of the lung cancer population receptive to a given targeted therapy increasingly small.

2.1. Patent Protection for Personalized Treatment

An additional issue with personalized medicine is the extent of patent protection of treatments. Because targeted therapies are based on very specialized molecules, and relatively small variations thereof may constitute new treatments without infringing on earlier patents,

[5] An exon is a segment of a DNA or RNA molecule containing information coding for a protein or peptide sequence.

[6] In some developing countries, Gefitinib and Erlotinib are still employed, as the third–generation inhibitors are not available.

the effective protection for targeted treatments is far shorter than the usual 20 years. In fact, new treatments for the same alteration may occur within a span of just a few years.

To illustrate, consider again Table 1. The first treatment for EGFR was discovered in 2004, and almost 10 years passed before an alternative was developed. However, targeted treatments for e.g. ALK alterations have been approved at much faster rate. Ceritinib (Novartis) was approved in 2014; just one year later, Alectinib (Roche) was approved. Subsequently, three additional treatments were approved in a span of four years: Crizotinib (Pfizer) in 2016, Lorlatinib (Pfizer) in 2018, and Brigatinib (Ariad) in 2020. Given the large investments in R&D that pharma companies have to bear (cf. the last column of Table 1), the short life span of these new treatments, and the increasingly smaller number of patients that they target, cast doubts on whether pharma companies will continue to have incentives to innovate. This is not just a hypothetical scenario: in February 2023, Roche announced that it was going to end the collaboration with Blueprint to further develop Pralsetinib, a drug for RET alteration (1%-2% of the lung cancer population) which is competing with Selpercatinib, a contemporaneous drug developed by Eli Lilly to treat the same alteration.

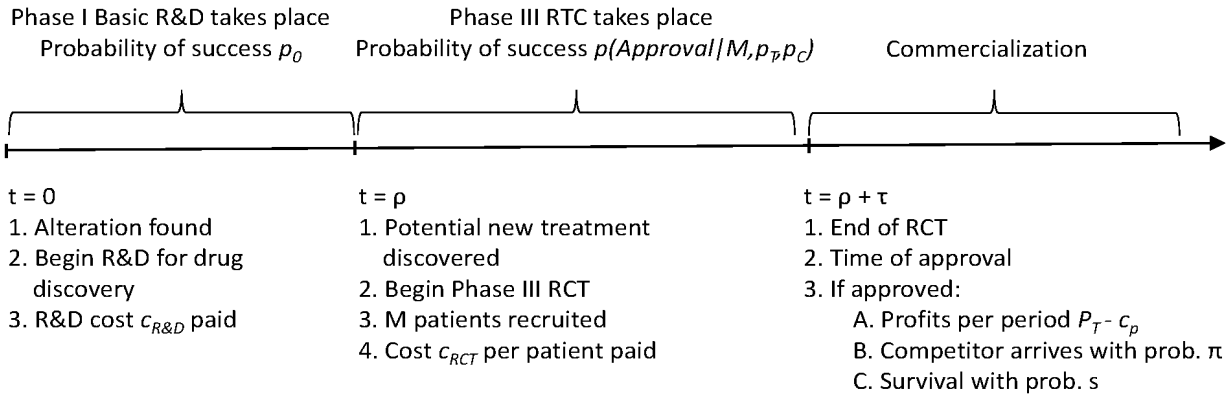
3. Model

Assume N new patients are diagnosed with a given illness per year. At time $t = 0$, a potential new alteration is discovered. The pharma company decides whether to develop a new drug targeting the alteration. Developing a treatment, testing it for safety and effectiveness in small-scale studies, and eventually obtaining approval from the regulatory authority after large-scale studies is a multi-stage, multi-year process: a description of the steps required by the US Food and Drug Administration (FDA) can be found at https://www.fda.gov/drugs/development-approval-process-drugs. For simplicity, we split the entire process into two phases, which, in a rough parallel with the FDA's terminology, we call "Phase I" and "Phase III." The first comprises identifying a new compound, testing it for safety in animals first and human volunteers later, and finally gathering preliminary data on its effectiveness;[7] we assume this stage entails a cost $c_{R\&D}$, takes ρ years to complete, and has success probability p_0.

The second phase comprises the large-scale randomized clinical trial (RCT) required to secure approval; we call this "Phase III," again in line with FDA's terminology. This is also a fairly complex process, which normally involves patient recruiting and coordination across multiple hospitals, complex protocols to adhere to, and so on. To capture the essence of

[7]This corresponds to the phases the FDA calls "animal testing," "IND Application," "Phase I," and "Phase II:" see https://www.fda.gov/drugs/information-consumers-and-patients-drugs/fdas-drug-review-process-ensuring-drugs-are-safe-and-effective.

Figure 4: Timeline of Drug Development



the decision, however, we simplify the process as follows. Assume the pharma company has identified M patients with the given alteration (we return to this point below). Conducting the RCT costs c_{RCT} per patient, net of any credit or savings, such as those provided by the Orphan Drug Act. The M patients are split into two groups of equal size, one treated with the new drug and the other with a control, which has a known success probability of p_C. (Success here can be measured in various ways, such as extending the survival probability by a given number of years; we provide further details in Section 4.) As for the new treatment, in reality, the pharma company is uncertain about its success probability p_T; we take this into account in our calibrations of Section 4.. For simplicity, in this section only, we assume that p_T is known to the pharma company. Overall, the RCT lasts τ years. Following approval, the pharma company commercializes the drug and earns a (monopoly) profit per patient for every year the patient is alive, until a competitor enters the market with a superior product. For simplicity, we assume that the new competitor completely displaces the pharma company's treatment; in reality, two or more drugs targeting the same alteration may coexist for a period of time (cf. Table 1). The overall timeline is depicted in Figure 4.

We first analyze the decision to carry out RCTs, conditional on having developed the drug, and having identified M patients with the alteration of interest. Then, we tackle the key issue introduced by personalized medicine—the fact that, out of the total pool of subjects recruited by a pharma company, only a fraction will in fact exhibit the alteration being studied. Finally, we work backward and consider whether pharma companies actually have incentives to invest in discovering a new treatment.

Let x_T and x_C be the number of individuals who are successfully treated in the treatment and control groups respectively. Conditional on M, the respective success rates are then

$$\widetilde{p}_T = \frac{x_T}{M/2}; \quad \widetilde{p}_C = \frac{x_C}{M/2}.$$

To bring about intuition, in this section we utilize a normal distribution approximation, as it is in fact regularly done in practice: e.g. Lo et al. (2022). Specifically, conditional on identifying M patients with the given alteration,

$$\widetilde{p}_T \approx N\left(p_T, \frac{p_T(1-p_T)}{M/2}\right); \quad \widetilde{p}_C \approx N\left(p_C, \frac{p_C(1-p_C)}{M/2}\right)$$

The trial is deemed successful if it clears the usual statistical bar for type 1 error:[8]

$$\widetilde{z} = \frac{\widetilde{p}_T - \widetilde{p}_C}{\sqrt{\frac{\widetilde{p}_T(1-\widetilde{p}_T)}{M/2} + \frac{\widetilde{p}_C(1-\widetilde{p}_C)}{M/2}}} > z(\alpha) \tag{1}$$

where, denoting by $\Phi(.)$ the standard normal cumulative density, $z(\alpha) = \Phi^{-1}(1-\alpha)$ is the threshold for a one-sided $\alpha-$confidence test under the null hypothesis $H_0 : p_T \leq p_C$. Conditional on success probabilities p_T and p_C, the probability of rejecting H_0 and thus accepting the new drug is

$$Pr(\text{Approval}_T | p_T, p_C, M) = 1 - \Phi\left(z(\alpha) - \kappa\sqrt{\frac{M}{2}}\right) \tag{2}$$

where $\Phi(.)$ is the standard normal distribution, and

$$\kappa = \frac{p_T - p_C}{\sqrt{p_T(1-p_T) + p_C(1-p_C)}} \tag{3}$$

is the standardized improvement of the treatment T against its control C.[9]

Clearly, ex-ante, as M declines, the probability of accepting the new drug (when $p_T > p_C$) declines. On the other hand, for given M, the probability of accepting the drug increases with the effectiveness of the new therapy compared to the earlier one, that is, as κ increases.

3.1. Pharma Profits and Discounting

Let P_T denote the expected price of the new treatment and c_p the marginal cost of producing it, once it has been approved. The profit to the pharma company per year and per individual is $(P_T - c_p)$ for as long as the pharma company remains a monopolist in the specific therapy, and zero when a better competitor appears and steals the market. Recall that we assume that drug development takes ρ years, and conducting an RCT takes τ years. Subject to approval, j periods after time $\rho + \tau$ (i.e., at calendar time $t = \rho + \tau + j$), there are $N \times$

[8] Clinical trials target not only the size of the test (α) but also its power (β). Sample sizes are normally chosen to match some values of both. For simplicity, we only consider the size of the test in this paper.

[9] The Bernoulli distribution of random variable Y has mean $E[Y] = p$ and variance $\sigma(Y)^2 = p(1-p)$. Hence, $\kappa = E[Y_T - Y_C]/\sigma(Y_T - Y_C)$.

$(1 + s + s^2 + .. + s^j) = N \times \frac{1-s^j}{1-s}$ patients to treat, where N is the number of new cases per year, and s is the (expected) progression free survival rate—that is, the probability that a patient is still alive and has not progressed to a more severe stage of the disease.[10] To account for competition, we assume that, in every period after $\rho + \tau$ there is probability π that a superior treatment is discovered. Thus, at each time $t = \rho + \tau + j$, the firm realizes a profit with "survival" probability $(1 - \pi)^j$. Finally, we let t_c be the corporate tax rate.

Given these assumptions, the present value of future profits, conditional on the drug being approved, is given by $A(P_T - c_p)N$, where A is the after-tax discount factor

$$A = (1 - t_c) \times \sum_{j=1}^{\infty} \frac{(1-\pi)^j}{(1+R)^{\tau+j}} \times \frac{1-s^j}{1-s} = \frac{(1-t_c)(1-\pi)}{(1+R)^{\tau}(R+\pi)} \times \frac{1+R}{1+R-(1-\pi)s} \tag{4}$$

and R is the pharma company's discount rate. We discount profits realized at dates $\rho + \tau + 1, \rho + \tau + 2, \ldots$ back to time $t = \rho$ because that is the point in time when the pharma company must decide whether or not to undertake the RCT.

3.2. Personalized Medicine

Personalized medicine allows physicians to start from the population of N patients with a given illness (e.g. cancer), perform additional tests (e.g. gene sequencing to detect alterations), and categorize patients into n mutually exclusive subgroups to receive targeted therapies. These tests are normally not expensive and one recently discovered test, the NGS (Next Generation Sequencing) test, can determine the specific genomic alteration (Mirza et al., 2024). We set this cost to be zero, for simplicity.

The key implication for our analysis is that, when conducting an RCT to test the effectiveness of a new drug against a specific alteration, the number M of patients with the given alteration is *not* a choice variable for the pharma company. To elaborate, let $i = 1, ..., n$ be an index denoting each subgroup. For simplicity, we assume here that each subgroup i has size N/n, although in reality some alterations are more common than others and groups of unequal sizes are the norm (see Figure 1). Given this simplifying assumption, a given patient belongs to a group $i = 1, .., n$ with probability $\frac{1}{n}$. If the pharma company recruits a total of K patients with the given illness to run a clinical trial for group i, the probability that M out of those K patients belong to group i is then

$$\Pr(M|K, n) = \binom{K}{M} \left(\frac{1}{n}\right)^M \left(1 - \left(\frac{1}{n}\right)\right)^{K-M} \tag{5}$$

[10] In the calibration of the next section, the survival rate is $s = s_T \times s_{other}$ where s_T is the survival rate due to the treatment, while s_{other} is the survival rate for other causes of death.

Hence, as the number of alterations n increases, the probability of securing a large enough sample M of patients to conduct an RCT for a *specific* subgroup declines. Indeed, since M follows a binomial distribution,

$$E[M|K,n] = \frac{K}{n} \tag{6}$$

As n increases, for a fixed total number K of subjects recruited, the expected number of patients in each subgroup declines. Thus, K must increase in order to allow for a sufficient number of patients in each RCT for each targeted therapy i. Clearly, K is bounded by N, the total number of patients with the given illness.

The choice variable for the pharma company is thus K. In addition, each new targeted therapy will only be beneficial to N/n patients, those in group i, rather than the full group of N patients. To sum up, as the number n of subgroups increases, it becomes harder to recruit a sufficient number of patients for each subgroup i, *and* the addressable market for each targeted therapy also shrinks.

Denoting P_i the market price of the drug in group i after successful approval, the present value formula for each treatment i when there are n groups is then

$$PV_{III,i}(n) = p_i(K,n)\, A\, (P_i - c_p) \left(\frac{N}{n}\right) - c_{RCT}\, E[M|K,n] \tag{7}$$

where A is as in Eq. (4) and $p_i(K,n)$ it the probability of approval, given by

$$p_i(K,n) = \sum_{m=1}^{K/2} \Pr[2m \leq M \leq 2m+1|K,n] \left[1 - \Phi\left(z(\alpha) - \kappa_i\sqrt{m}\right)\right]. \tag{8}$$

where $\kappa_i = \frac{p_i - p_C}{\sqrt{p_i(1-p_i)+p_C(1-p_C)}}$. To interpret, suppose zero or one out of K recruited patients exhibits alteration i: then, the RCT cannot be run, and so the new treatment cannot be approved. If $2m \geq 2$ patients exhibit the alteration, then m are assigned to the treatment group, and m to the control group. The same occurs if $2m+1$ patients exhibit alteration i; the $(2m+1)$-st patient is not assigned to either group.

In Eq. (7), as n increases, the present value of treatment i decreases mechanically, because the number of patients in group i decreases. However, it may still be the case that, if we take the perspective of the entire pharma industry and aggregate over all n groups, the total present value of the industry increases, or is at least sufficient to induce investment in personalized treatments. We now turn to this question and show that, in fact, even the present value of the pharma industry as a whole eventually decreases as n grows large. We assume that the prices of all targeted therapies are the same: $P_i = P_{TT}$. Furthermore, we assume that the likelihood of success is the same for all targeted therapies: $p_i = p_{TT}$, so $\kappa_i = \kappa$. Then, aggregating across all n groups, we have

$$PV_{III}(K,n) \;\; = \;\; p(K,n)\, A\, (P_{TT} - c_p)\, N - c_{RCT}\, K \tag{9}$$

where $p(K,n)$ is as in Equation (8) with $\kappa_i = \kappa$.

We can view non-personalized medicine as the special case in which $n = 1$. In that case, (9) shows the present value of conducting a Phase-III RCT when the number of subjects with the given illness is itself a choice variable—i.e., $M = K$.

To study the comparative statics of the choice of optimal sample size K, we make the simplifying assumption that, in calculating present values, the pharma company approximates both the distribution of M and the approval probability for a given M by suitable normal distributions. As noted above, at least the latter approximation is common practice in the pharma industry.[11] We then have the following assumption:

Assumption (Normal Approximation). The pharma company calculates $PV_{III}(K,n)$ assuming that

$$M \sim \phi(M, K/n, K/n(1-1/n)) \quad \text{and} \quad \Pr(\tilde{z} > z(\alpha)|M) = 1-\Phi\left(z(\alpha) - \kappa\sqrt{M/2}\right) \text{ for } M \geq 0$$

where $\phi(.,a,b)$ is the normal density function with mean a and variance b, and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Proposition 1 *Under the Normal Approximation Assumption, the probability of approval $p(K,n)$ is increasing in K:*

$$\frac{\partial p(K,n)}{\partial K} > 0 \tag{10}$$

That is, as K increases, it is increasingly more likely that the clinical trial will secure a large number M of patients, and therefore that the FDA will approve the new treatment.

However, a higher K increases the costs of clinical trials. As K increases, the second term in Eq. (9) may become so large as to overcome the benefit from increasing the total patient pool K, and $PV_{III}(n)$ may decrease in K. Unfortunately, $PV_{III}(n)$ is not a concave function of K in general. However, if we explicitly assume that it is (as is the case in the relevant region in our simulations), we obtain

Proposition 2 *Under the normal approximation assumption, let $PV_{III}(K,n)$ be concave in K. Everything else equal, the optimal number of patients K^* is (1) increasing in profits per period $(P_{TT} - c_p)\,N$; (2) decreasing in the discount rate R; (3) decreasing in the probability of a competing treatment emerging, π; (4) increasing in the survival rate of the new treatment s; (5) decreasing in time to production τ; (6) decreasing in the cost per person of Randomized Clinical Trials c_{RCT}.*

[11] In our calibrations in Sec. 4., we also verified that the normal approximation is acceptable for the parameter values we consider.

In addition, if the density $\phi(M, K/n, K/n(1-1/n))$ *is such that* $z(\alpha) < \kappa\sqrt{M/2} - \frac{1}{\sqrt{M/2}}$ *is satisfied with sufficiently high probability,*[12] *then the optimal number of patients* K^* *is decreasing in* κ*.*

This proposition shows that our model generates natural comparative statics results; it will help interpret our calibrated results below. However, in the area of personalized medicine, the critical parameter is n, the number of possible alterations that targeted therapies aim to affect. We discuss next the impact of n on optimal number of patients required in randomized clinical trials, and thus the incentives to even develop new drugs.

3.3. R&D Disincentives as the Number of Targets n Increase

For our next results, we maintain the Normal Approximation assumption, with the further assumption that n is sufficiently large so that the approximate variance of M, namely $K/n(1-1/n)$, is approximately K/n:

Assumption (Normal Approximation) 2. In calculating $p(K, n)$, the pharma company adopts the approximation

$$M \sim \phi(M, K/n, K/n). \tag{11}$$

Under this assumption, K and $1/n$ enter symmetrically in the probability of approval. The following result then follows from a minor modification of the proof of Proposition 1; we emphasize it because it is critical for the argument that follows:

Proposition 3 *Let the Normal Approximation Assumption 2 hold. Then, probability of approval* $p(K, n)$ *is decreasing in* n*:*

$$\frac{\partial p(K, n)}{\partial n} < 0 \tag{12}$$

That is, as the medical profession identifies more and more targets for targeted therapies, for given number of patients K in randomized clinical trials, it becomes increasingly less likely that these new therapies will be approved. This implies that the present value of future profits decline, everything equal:

Proposition 4 *Let the Normal Approximation Assumption 2 hold. Then, everything else equal, the present value* $PV_{III}(K, n)$ *decreases as* n *increases.*

[12] The exact condition is that $cov\left(M^2, \phi\left(z(\alpha) - k\sqrt{M/2}\right)\sqrt{M/2}\right) < 0$, which is satisfied when $\phi\left(z(\alpha) - k\sqrt{M/2}\right)\sqrt{M/2}$ decreases in M, i.e. when the condition in the text is satisfied.

The previous results show that as the number of alterations n increases, the present value of future profits declines as the probability of approval $p(K, n)$ declines. These results hold locally under the normal approximation assumption. We can also show a general limiting result for n large which does not depend on the normal approximation assumption:

Proposition 5 *For every value of A, P_T, c_p and N, there is $\bar{n}$ such that $PV_{III}(K, n) < 0$ for all $n \geq \bar{n}$.*

Proposition 5 captures a straightforward, yet important insight: if every patient is truly unique (which is essentially the case if n is large relative to N), then it will be virtually impossible for pharma companies to conduct an RCT, simply because, for any alteration, there will be a single patient in the treatment group, and *no* patient that can serve as control. Under the current approval process in which regulatory agencies approve individual therapies, a truly personalized medicine is economically infeasible.

Our results assume that κ is constant as n increases. Of course, the promise of personalized medicine is that, over time, new targetable alterations will be discovered, leading to more effective treatments. In other words, over time, both n and κ should increase. However, this assumes that therapeutic improvements are measured against the same baseline—say, chemotherapy. In reality, at any point in time, new treatments are tested against the current standard of care—which, precisely due to the development of personalized medicine, may itself be a previous targeted treatment. For instance, from Table 1, the first EGFR TKI Gefitinib was tested against chemotherapy, but later TKIs such as Dacomitinib and Osimertinib were tested against Gefitinib as control. This implies that the effective n and κ need not be correlated over time. Appendix A1. presents a concrete illustration of this issue.

3.4. Solving Backward: The Ex-ante Incentive for Personalized Drug Development

The results in the previous section show that even after having developed a new treatment, as the number of alterations n increase, the probability of approval declines and so does the present value of future profits. We now analyze the impact of the approval process on the ex-ante decision to develop the drug in the first place—that is, in Phase I.

As per Fig. 4, the first step in the drug development process is to identify a treatment for a specific alteration i, and then to carry out the initial testing. Let $c_{R\&D,i}$ denote the R&D cost incurred in this phase. Let $p_{0,i}$ be the probability of advancing to Phase III (i.e. to the randomized clinical trial), and denote by $A_0 = (1 + R)^{-\rho}$ the discount factor corresponding to Phase I, which takes into account the time ρ between Phase I and Phase III.[13] Then, the

[13]For simplicity, we do not take into account the probability that a competitor will develop a treatment

present value of carrying out the R&D for the new drug for alteration i, and subsequently conducting an RCT to secure approval for it, is

$$PV_{I,i} = p_{0,i}\, A_0\, PV_{III,i} - c_{R\&D,i} \tag{13}$$

As before, we assume symmetry and add up across groups to obtain the (industry-wide) present value of developing drugs for all alterations:

$$PV_I = p_0\, A_0\, \{p(K,n)\, A\, (P_{TT} - c_p)\, N - c_{RCT}\, K\} - c_{R\&D}\, n \tag{14}$$

At time 0, i.e., at the beginning of Phase I, there are two reasons why expected profits decline as the number of alternations n increase. The first is that, as before, conditional on developing a drug for a given alteration, expected profits decline due to the decreased probability of approval $p(K,n)$. In addition, from the ex-ante perspective, the cost of R&D increases proportionally with n: for each new drug, the pharma industry as a whole has to support an expensive development process. Even if $c_{R\&D}$ were to decline significantly due to e.g. technological change or learning, the structure of randomized control trials still implies that PV_I is still likely to decline due to the effect discussed in previous sections. Indeed, this true even if $c_{R\&D} = 0$: in this case, the present value in Eq. (14) only differs from the one in Eq. (9) by a factor $p_0 A_0$.

The same forces also adversely affect the break-even price P_{TT} that makes $PV_I = 0$:

$$P^{BE}_{I,TT} = c_p + \frac{p_0\, A_0\, c_{RCT}\, K + c_{R\&D}\, n}{p_0\, A_0\, p(K,n)\, A\, N}. \tag{15}$$

Proposition 6 *The break-even price $P^{BE}_{I,TT}$ is increasing in the number of alterations n. Moreover, for every value of $A, c_p, c_{R\&D}$ and N, and for every $\overline{P}$, there is $\overline{n}$ such that $P^{BE}_{I,TT} > \overline{P}$ for $n > \overline{n}$.*

Proposition 6 shows that the break-even price becomes unboundedly large as n increases. This also implies that societal costs for targeted therapies is unbounded. Suppose the social planner operates the pharma company directly for the purpose of developing, producing and selling the targeted therapy for the lowest feasible annual price per patient. The latter price should be just enough so that expected discounted profits exactly cover R&D and RCT costs. This corresponds to the break-even price in Eq. (15) when both the corporate tax rate t_c and tax credits affecting c_{RCT} (see Section 3.) are set to zero. Since patients' survival rate is s_T, the *total* annual societal cost can then be calculated as $P^{BE}_{I,TT} \times N/(1 - s_T)$ where $N/(1 - s_T)$ is the average number of patients treated per year in steady state.

for the same alteration. This makes our analysis more conservative, as the present value at time 0 would be lower if we did.

Unbounded societal costs occur even if the $R\&D$ cost of Phase I, $c_{R\&D}$, is zero: the driving factor is the decline in the probability of approval $p(K, n)$. Hence, even if technological advances such as machine learning and AI substantially decrease the costs of early-phase R&D, personalized medicine will still be unsustainable under the current approval regime.

One caveat is that there may be spillover effects across groups, so that the cost $c_{R\&D}$ may be a function of n itself. Indeed, we expect that $c_{R\&D}$ for targeted therapies will decline substantially in the future as n increases. Still, unless $c_{R\&D}$ decreases faster than $\frac{1}{n}$, the overall effect of an increase in n will still be to increase the break-even price.

4. Calibration to Lung Cancer Research

In this section we provide a calibration which illustrates the issues surrounding personalized medicine. We use parameters from lung cancer research. We present two sets of results, one for the U.S. only, and one for the entire world.

4.1. Parameters

Table 2 collects the parameters we use in the two calibrations. First, from Figure 1, about 50% of total lung cancer patients do not have targetable alterations and so they are not candidates for any targeted therapy. In addition, about 50% of lung cancer patients are currently not tested at all (see e.g. Vidal et al., 2023) due to lack of awareness or technology in many hospitals. There are about 220,000 new lung cancer cases per year in the US.[14] Thus, we set the number of new targetable cases per year at $N = 220,000/4 = 55,000$.

We set the arrival probability of a superior competitor at $\pi = 1/3$. That is, in expectation, it takes three years before a superior competitor arrives on the market. This probability is roughly in line with the data in Table 1, which shows an acceleration in the discovery of new drugs for each alteration.[15] The length of a clinical trial is set at $\tau = 3$ years.

As for the success rate of the new treatment, Table 1 shows that the median progression-free survival (mPFS) varies greatly depending on the alteration and therapy. Also recall from Table 1 that the control arm is not necessarily chemotherapy, but rather the best previously approved drug. For instance, for EGFR, Osimertinib appears to be the best treatment (mPFS = 18.9 months) but it is only 8.7 months better than Gefitinib. For ALK, on the other hand, Alectinib is far better than Crizotinib, and Lorlatinib far better yet (NR means "Not Reached" in that more than 50% of patients are still progression free even after so

[14] See e.g. https://www.cancer.org/cancer/types/lung-cancer/about/key-statistics.html.

[15] The data in Table 1 also indicates that many therapies for the same alteration may be available at the same time, in which case firms must split the market.

many years). However, there is significant variation across treatments. For illustration, we use Osimentirib for the treatment group and Gefitinib for the control group.

To translate the mPFS into an annual frequency, recall that mPFS is the number of months during which 50% of patients are still progression-free. Assuming a constant annual survival rate s_T for the treatment group, this implies that $s_T^{mPFS} = 0.5$ and thus $s_T = 0.5^{1/(mPFS/12)}$. Since we assume that the targeted length of the RCT is $\tau = 3$ years, we obtain $p_T = s_T^{\tau} = 0.5^{\frac{\tau}{mPFS/12}}$. These calculations give a survival rate and success probabilities of $s_T = 0.644$ and $p_T = 0.2671$, and $s_C = 0.4424$ and $p_C = 0.0866$ for Osimertinib (treatment) and Gefitinib (control) respectively. In addition, we multiply these survival rates by the survival rate reflecting "other" causes of death unrelated to cancer, which we set to $s_{other} = 0.95$, for an average life expectancy of 20 years after diagnosis.

We also introduce ex-ante uncertainty from the perspective of pharma companies, as there is no guarantee that the new treatment will certainly be better than the control. We thus assume that p_T is a truncated normal distribution with $\tilde{p}_T \sim TrN(\mu_{P_T}, \sigma^2_{p_T}, 0, 1)$. We set $\sigma_{P_T} = (p_T - p_C)/0.5244$: that is, there is approximately a 30% chance that the new treatment is not better than the old treatment. This is optimistic, since about 36% of clinical trials with biomarkers in oncology fail in Phase III (see Wong et al., 2019, Table 3). The mean of the normal distribution μ_{p_T} is chosen so that $E[\tilde{p}_T] = p_T$. Similarly, Wong et al. (2019), Table 3, shows that the probability of reaching Phase III with biomarkers in oncology is just $p_0 = 0.168$; we adopt this number in our calibration.[16] Finally, we set the R&D cost pre-Phase III to between \$50 and \$500 million. These are conservative numbers compared to the costs calculated from Henderson et al. (2023, Supplemental Table S1) and reported in Table 1. We chose a range of lower values in part to take into account potential technological improvements that may lower the costs of R&D in the future.

Turning to the discount rate R, based on DiMasi, Grabowski, and Hansen (2016), we set $R = 10.5\%$. We let the corporate tax rate be $t_c = 21\%$, and reduce the cost of running RCTs (c_{RCT}) by $tc_{OD} = 25\%$ tax credit as provided by the Orphan Drug Act in the US.[17] Plugging the above values in Eq. (4), we find $A = 1.4115$ in the present-value formula (9).

Finally, as shown in Table 1, the price of a new drug varies greatly among targeted therapies. Using EGFR drugs as an example, we set the annual price of the new treatment at $P_{TT} = \$4000 \times 52 = \$208,000$. This is roughly in line with Skinner et al. (2018) for the annual cost of targeted therapies. We assume the marginal cost of producing the drug c_p is zero (equivalently that $P_{TT} = \$208,000$ is the net price per year): Leighl et al. (2021) report

[16] We also ran the calibration with the far higher value $p_0 = 0.45$, obtaining qualitatively similar results.

[17] We implicitly assume that the tax credit offsets the tax liability of pharma companies from other business lines. The qualification as an Orphan Drug may also provide market exclusivity for seven years, but this benefit does not seem to accrue to targeted therapies due to the specifics of the molecular alterations.

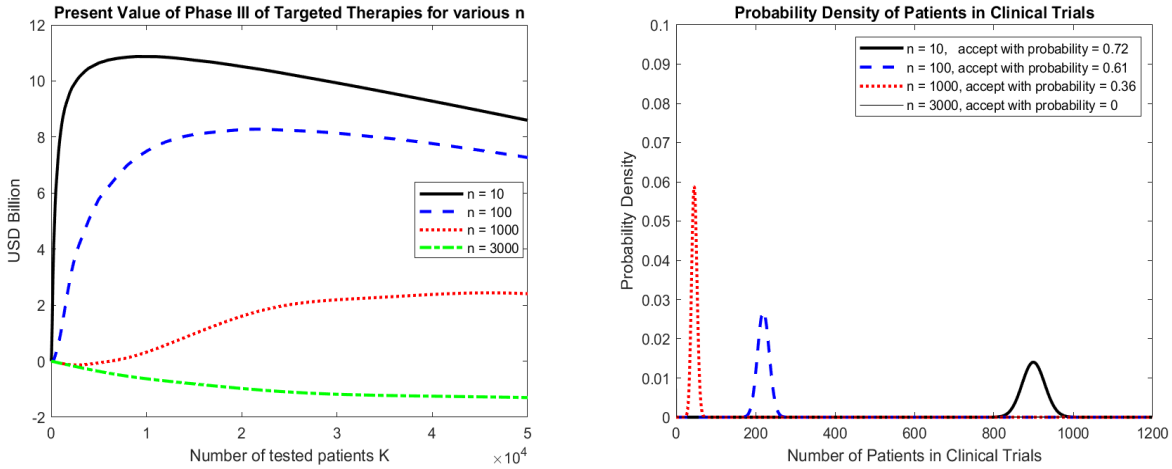
Table 2: Parameters in Calibration

Description	Parameter	value
U.S.		
New cases per year	$\overline{N}$	220,000
Testable New Cases	adj_1	0.5
Average Percent of Tested Patients	adj_2	0.5
Actual new cases per year	N	55,000
Pharma's discount rate	R	10.5%
Arrival probability of better competitors	π	$0.\overline{3}$
Length of clinical trial	τ	3 years
Median Progression Free Survival for T	$mPFS_T$	18.9 months
Median Progression Free Survival for C	$mPFS_C$	10.2 months
Annual survival rate for T	s_T	0.644
Annual survival rate for C	s_C	0.4424
Annual survival rate for other causes	s_{other}	0.95
Success probability at τ for T	$E[p_T]$	0.2671
Dispersion of p_T	$\sigma[p_T]$	0.2148
Success probability at τ for C	p_C	0.0866
Corporate tax rate	t_c	21%
Orphan Drug tax credit	tc_{OD}	25%
Resulting discount factor	A	1.4115
Annual net price of new treatment per patient	$P_{TT} - c_p$	$\$208,000$
Annual cost of RTC per patient	c_{RTC}	$\$100,000$
Success probability of Phase I	p_0	16.8%
World		
New cases per year	$\overline{N}$	2,500,000
Actual new cases per year	N	625,000
Adjustment for lower prices & lower patients	x	0.196
Adjusted net price of new treatment per patient	$P_{TT} - c_p$	$\$40,768$

that the manufacturing costs for targeted therapies are between 0.2% and 3% of the drug's price tag. The cost of a clinical trial is set at $c_{RCT} = \$100,000$ per patient (see Moore et al., 2020, Table 3 and Sertkaya et al. (2024), supplementary table 1).

The above parameter values are for the US market only. The bottom panel of Table 2 shows the modified parameters for the worldwide market. We first increase the total number of patients to $\overline{N} = 2.5$ million worldwide. The annual net price of new treatments is lower, as targeted therapies are sold at lower prices outside the US (see e.g. Goldstein et al. (2017)). Furthermore, targeted therapies are not as broadly available in many foreign countries as they are in the US, due to stricter regulatory requirements, and/or high costs. To jointly capture these price and market penetration differences, we leverage the fact that U.S. spends about 45% of the total global spending in cancer drugs (see IQVIA, 2024, exhibit 53.) This

Figure 5: **Phase III: Present Value and Probability Density**



implies that $45\% = \frac{P_{TT} \times N^{US}}{P_{TT} \times x \times N^{W}}$ where N^{US} and N^{W} are the US and world population of lung cancer patients respectively, and x is the adjustment factor capturing price and market penetration differences. Since the number of U.S. lung cancer patients N^{US} is about 8.8% of all lung cancer patients worldwide ($N^{US} = 0.088 \cdot N^{W}$), we conclude that $x = .196$. We thus assume that the worldwide price of targeted therapies is 19.6% of the US price.

The appendix describes the Monte Carlo simulations that take advantage of small sample statistics to calculate the various quantities in formulas (9), (14) and (15).

4.2. Results: Phase III, U.S. Calibration

The two panels Figure 5 illustrate the main issue with the increase in the number of potential alterations n, and the corresponding decrease in the probability $(1/n)$ of each individual belonging to a given group. The left panel plots the Phase-III present value of future profits from Eq. (9) against the number of tested individuals, for $n = 10$, 100, 1000 and 3000. Recall that the present value is for a Phase III trial for the pharma industry as a whole—that is, including all alterations. This allows us to compare the present value across different numbers of alterations n while keeping constant the total new number N of patients per year. Consider first the solid black line, which corresponds to $n = 10$, and hence a frequency of 10%. This line increases as K increases, due to the increase in probability of acceptance from increasing the number of patients in the clinical trial; then it peaks at around $K = 9,000$ patients tested, and subsequently declines steadily. The black line on the right panel shows the probability density of patients in clinical trials M when we test $K = 9,000$ subjects, and there are $n = 10$ alterations. The distribution is concentrated around $M = 900$ patients in each clinical trial, i.e. one out of 10. With that large number of patients, the total probability

that the therapy is accepted is 72%, as reported in the legend.

Consider now the blue dashed line in the left panel, which corresponds to $n = 100$. In this case, the corresponding frequency of each alteration is 1%, which is still higher than many alterations shown in Figure 1. The optimal number of patients is now $K = 22,000$, and the present value in Phase III is uniformly lower than the case of $n = 10$ alterations. The dashed blue line in the right panel shows that the corresponding distribution of patients in clinical trials now concentrates around $M = 220$, i.e. one out 100. The probability of accepting each new drug is now 61%.

When the number of groups for target therapies increase further to $n = 1000$ (dashed red line), the optimal number of patients increases to $K = 50,000$. The dashed red line shows a much smaller present value than previous case, as the probability of acceptance also declines substantially. The right panel shows that the density of patients in clinical trials now concentrates around 50 per trial, i.e. one in $1,000$. At this low number of patients, the probability of acceptance drops to just 36%.

Further increasing the number of alterations to $n = 3,000$ makes the present value PV_{III} negative, and so targeted therapies are no longer viable. With $3,000$ alterations and $62,500$ total patients available for clinical trials, even enrolling *all* patients in RCTs, each treatment would get only an average of 20 patients, which is too small to guarantee statistical significance. Moreover, the large costs of running clinical trials on all new cases make the present value negative (see Proposition 5).

The left panel of Figure 6 shows the optimal number of patients tested K and the corresponding optimal present value as n increases from 1 to $10,000$. For each n, the graph shows the present value PV_{III} in Eq. (9) against K, and the solid black line represents the combination (K^{max}, PV_{TT}^{max}) for each n. As it can be seen, the maximal present value declines as n increases and the optimal number of patients tested K^{max} mostly increases. Beyond a certain number n of alterations, however, K^{max} declines, eventually dropping to $K^{max} = 0$, meaning that it is best not to pursue clinical trials when n is this large.

The right panel of Figure 6 shows the corresponding pairs of optimal subject pool sizes $K^{\max}$ and acceptance probabilities, when n ranges from 1 to $10,000$. As it can be seen, the probability of acceptance of new therapies initially declines mildly as the optimal K^{max} increases. The decline in the present value is due to the increasing cost of a large population in the randomized control trials. However, at some point, the probability of acceptance of the new therapy decreases substantially, making it optimal to actually decrease the number of patients tested K, moving rapidly towards zero.

Finally, Figure 7 plots the maximized present value in Phase III against the number of alternations n. This figure illustrates that the present value of conducting clinical trials in

Figure 6: **Present Value of Phase III and Acceptance Probability across n and K**

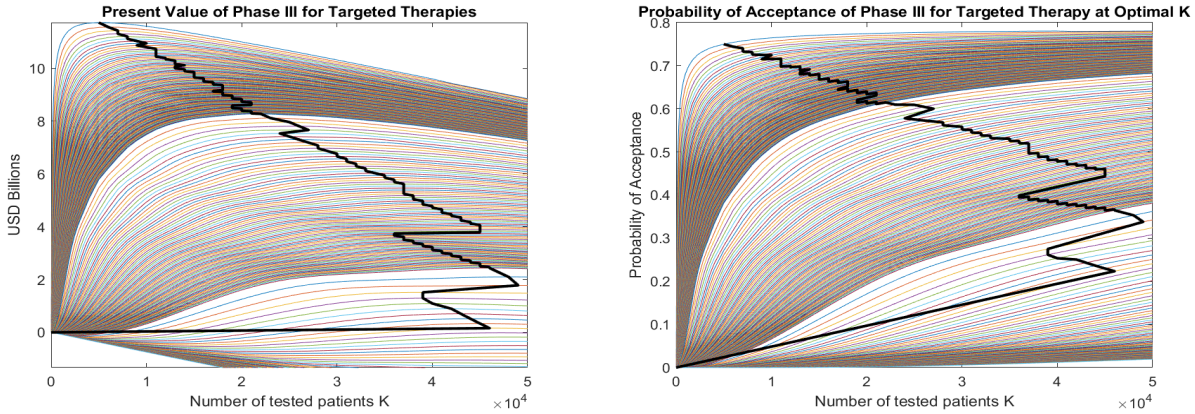
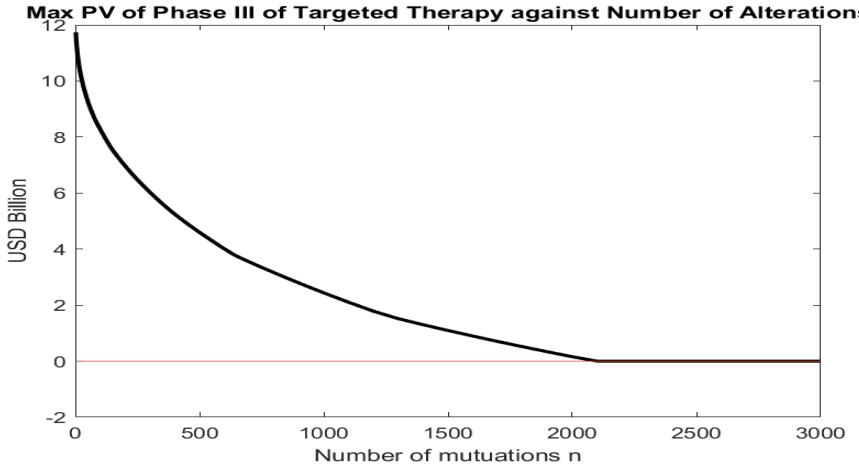


Figure 7: **Maximized Present Value of Targeted Therapies in Phase III**



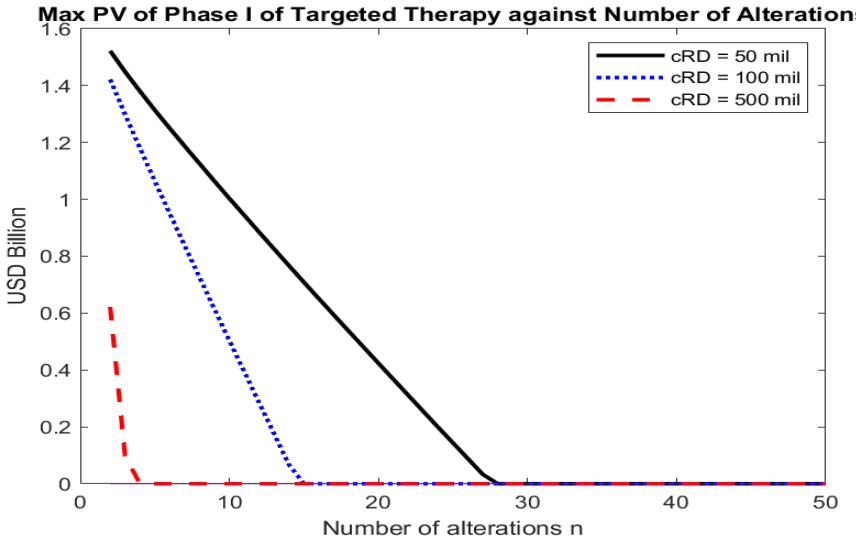
the pharma industry as a whole drops quickly as n increases, reaching zero for just over $n = 2,000$, i.e., for an incidence of just below 0.05%. That is, personalized medicine is not viable in the current system of clinical trials, even when potential targeted therapies are available (i.e. we reached Phase III).

4.3. Results: Phase I, U.S. Calibration

Moving backwards, pharma companies who anticipate the higher costs of randomized control trials and the decreasing likelihood of acceptance of the new therapies under the current approval regime will eventually not invest in new therapies to start with. Table 1 shows the ex-post R&D costs for those drugs that make it to the market.

Figure 8 shows the maximized present value as of Phase I for three levels of costs $c_{R\&D} =$ \$50, \$100, and \$500 millions. As can be seen, even for a relatively small R&D costs of R&D

Figure 8: **Max Present Value of Future Profits in Phase I, US Calibration**



of \$50 million, the present value from investing in drug development drops to zero for n less than 30—that is, for an incidence of alterations of over 3.3%.

Pharma companies can increase prices of targeted therapies to cover their R&D costs. Panels A and B of Table 3 show the break-even price per year per patient, and the societal costs, respectively. For these calculations, we set the corporate tax t_c and the Orphan Drug tax credit to zero to estimate societal costs (see discussion after Proposition 6). We consider five different levels of $R\&D$ costs, spanning from \$5 million to \$500 million. The latter figure is in line with the current ex-post costs of R&D as per Table 1. Break-even prices become prohibitive when n increases. Even for a relatively low number of alterations $n = 200$, which implies that each alteration affects 0.5% of the lung cancer population, the minimum break-even price for R&D costs per drug of just $c_{R\&D} = \$50$ million is $P^{BE}_{I,TT} = \$1.3$ million per year, per patient. Assuming more realistic R&D costs of \$500 million per drug, the break-even price rises to $P^{BE}_{I,TT} = \$11.9$ million per year, per patient. When $n = 2,000$, for an incidence of 0.05% in the lung cancer population, the break-even price equals $P^{BE}_{I,TT} = \$31.2$ and \$310 million per year, per patient, for $c_{R\&D} = \$50$ and $= \$500$ million, respectively.

At these break-even prices, societal costs are enormous. We collect them in Panel B. Consider the case with a number of alterations $n = 2,000$ (0.05% of the lung cancer population per alteration). Recall that this is realistic, in view of the data in Figure 1. In this case, societal costs would be \$4.8 trillion per year when $c_{R\&D} = \$50$ million, and \$48 trillion per year when $c_{R\&D} = \$500$ million.

We can also assess the magnitude of societal costs by comparing our annual break-even costs of developing targeted therapies to the estimates Value of Statistical Life Year

(VSLY) from the U.S. Department of Health and Human Services (HSS).[18] The Value of Statistical Life estimates the amount individuals are willing to pay to marginally increase life expectancy. VSLY is its annualization which takes individuals' life expectancy into account. HHS benchmark estimates of VSLY range between $231,000 and $754,000. A glance at Panel A of Table 3 shows that annual break-even costs of developing targeted therapies are far higher than these values once the number of alterations n is over 500, i.e. for an incidence of 0.1%.

4.4. Results: Phase I, World Calibration

We now discuss the results when we include the rest of the world. This case is more complicated, as multiple regulatory agencies must approve the therapies, which could happen in a staggered fashion, and at higher costs. Here, we assume that if one regulatory agency approves a treatment, so do the others—an optimistic assumption. In this case, the number of patients to be treated would increase substantially, as world-wide, around 2.5 million people are diagnosed with lung cancer every year. However, as noted above, drug prices around the world are far lower than in the US, and targeted treatments are less common.

As discussed in previous section, we adjust the price P_{TT} down for the rest of the world by a factor of 0.196, which matches the fact that U.S. cancer drug spending is about 45% of global spending. Again, the adjustment takes into account a mix of lower prices and/or lower fraction of the population treated with targeted therapies outside the U.S.

Figure 9 shows the maximum present value of future profits as of Phase I for three levels of R&D costs. For current expected costs of development of $500 million, this present value would reach zero at less than $n = 10$ alterations. If the cost of $R\&D$ were to drop to $50 million, the present value would reach zero at $n = 60$ alterations. Again, as n increases, the incentive of the pharma industry as a whole to undertake research in targeted therapies declines substantially.

Panels C and D of Table 3 show the break-even price per year per patient, and the societal costs, respectively, for the world calibration. The break-even price from formula (15) needs to be divided by the factor $(.088 + (1 - .088)\,0.196)$ to take into account the lower prices and/or lower fraction of the population treated outside the U.S., where 0.088 is the fraction of U.S. lung cancer patients in the world total, and 0.196 is the adjustment required to ensure that U.S. drug spending equals 45% of the total.[19] As in previous section,

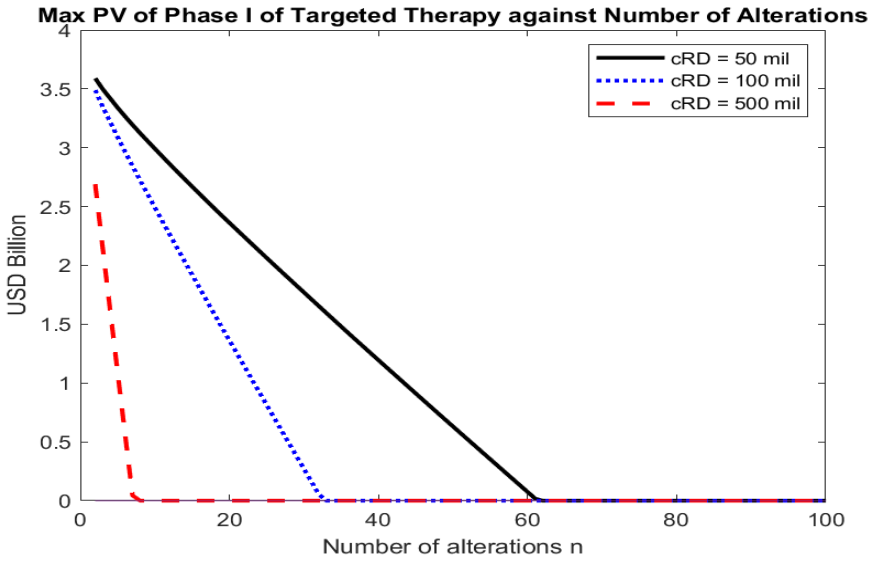
[18] See e.g. https://aspe.hhs.gov/sites/default/files/documents/cd2a1348ea0777b1aa918089e4965b8c/standard-ria-values.pdf

[19] To see this, the present value formula is $PV_{I,TT} = A_0 p_0 [p(K,n) A (N^{US}(P_{TT}^{US} - c_p) + (N - N^{US})\, x\, (P_{TT}^{US} - c_p)) - K c_{RCT}] - n c_{R\&D}$ which gives $P_{TT}^{US} = c_p + \frac{A_0 p_0 K c_{RCT} + n c_{R\&D}}{A_0 p_0 p(K,n) A N (.088 + (1 - .088)\, 0.196)}$

Table 3: Phase I Break-even Prices and Societal Costs of Targeted Therapies

Panel A: Breakeven Annual Cost per Patient – U.S. Calibration						
Number n	Percent in	Cost of R&D				
of alterations	Population	5 million	10 million	50 million	100 million	500 million
10	10	9,146	15,642	61,511	115,091	522,188
100	1	91,483	156,433	614,743	1,156,884	5,485,270
200	0.5	182,939	312,853	1,264,313	2,446,893	11,907,531
500	0.2	463,853	827,244	3,734,371	7,368,281	36,439,556
1000	0.1	1,111,397	2,087,824	9,899,243	19,663,516	97,777,703
2000	0.05	3,318,447	6,422,370	31,253,751	62,292,977	310,606,780
5000	0.02	29,496,301	58,199,095	287,821,450	574,849,390	2,871,072,900
10000	0.01	199,992,700	397,258,640	1,975,386,100	3,948,045,400	19,729,320,000
Panel B: Breakeven Societal Cost (million of USD)– U.S. Calibration						
Number n	Percent in	Cost of R&D				
of alterations	Population	5 million	10 million	50 million	100 million	500 million
10	10	1,413	2,417	9,503	17,781	80,675
100	1	14,134	24,168	94,974	178,732	847,443
200	0.5	28,263	48,334	195,329	378,031	1,839,647
500	0.2	71,663	127,804	576,939	1,138,358	5,629,707
1000	0.1	171,705	322,557	1,529,377	3,037,903	15,106,106
2000	0.05	512,681	992,220	4,828,529	9,623,915	47,987,003
5000	0.02	4,557,013	8,991,433	44,466,797	88,811,001	443,564,633
10000	0.01	30,897,749	61,374,228	305,186,055	609,950,834	3,048,069,101
Panel C: Breakeven Annual Cost per Patient – World Calibration						
Number n	Percent in	Cost of R&D				
of alterations	Population	5 million	10 million	50 million	100 million	500 million
10	10	3,015	5,150	20,258	37,999	172,340
100	1	30,153	51,505	202,540	379,986	1,723,407
200	0.5	60,300	103,010	405,166	759,739	3,446,816
500	0.2	150,857	257,615	1,012,700	1,899,223	8,640,834
1000	0.1	301,539	515,052	2,025,401	3,798,780	17,794,450
2000	0.05	603,005	1,030,105	4,095,260	7,867,255	38,043,212
5000	0.02	1,507,613	2,630,347	11,612,216	22,839,552	112,658,241
10000	0.01	3,411,917	6,324,593	29,625,999	58,752,759	291,766,821
Panel D: Breakeven Societal Cost (million of USD)– World Calibration						
Number n	Percent in	Cost of R&D				
of alterations	Population	5 million	10 million	50 million	100 million	500 million
10	10	1,407	2,404	9,455	17,734	80,433
100	1	14,073	24,038	94,528	177,345	804,337
200	0.5	28,143	48,076	189,097	354,581	1,608,675
500	0.2	70,407	120,232	472,641	886,393	4,032,794
1000	0.1	140,732	240,382	945,282	1,772,942	8,304,910
2000	0.05	281,430	480,764	1,911,313	3,671,754	17,755,280
5000	0.02	703,623	1,227,618	5,419,578	10,659,527	52,579,120
10000	0.01	1,592,388	2,951,773	13,826,853	27,420,705	136,171,509

Figure 9: **Max Present Value of Future Profits as of Phase I, World Calibration**



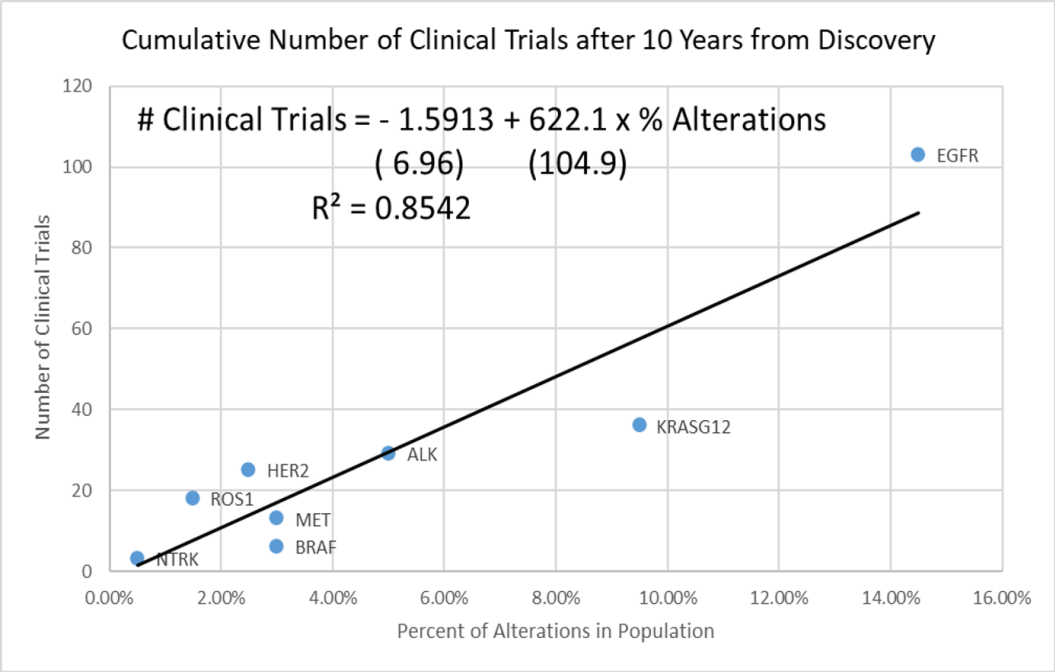
we set the corporate tax t_c and the Orphan Drug tax credit to zero to estimate societal costs. Thanks to a larger population to treat, the average prices per patient are lower than in the US calibration (Panel A), but they are still substantial. For $n = 200$, for instance, the break-even price is $P^{BE}_{I,TT} = \$551,423$ per year per patient when the cost of R&D is $c_{R\&D} =$ \$50 million per drug, and $P^{BE}_{I,TT} =$ \$4.7 million per year per patient when the cost of R&D is \$500 million per drug.

The societal costs in Panel D are again prohibitive. We calculate the societal costs by multiplying the break-even price by the coefficient $N(.088+(1-.088)\,0.196)/(1-s_T)$ again to take into account the lower prices and lower population treated outside U.S. For $n = 2,000$, for instance, societal costs are \$1.9 trillion per year for R&D costs $c_{R\&D} =$ \$50 million per drug, and \$17.7 trillion per year for R&D costs $c_{R\&D} =$ \$500 million per drug. As n increases, the global societal costs keep increasing, all the way to \$136 trillion when $n = 10,000$ per year if R& D costs are \$500 million per drug.

4.5. Personalized Medicine: All Diseases as Rare Diseases

The key conclusion we draw from these calibrations is that the discovery of alterations that open the door to personalized targeted therapies are effectively transforming a common disease like lung cancer, which affects over 220,000 people in the US per year, into a collection of rare diseases. Our results show that, due to economic incentives, such alterations may become under-studied by both researchers and the pharma industry—as indeed is the case for other rare diseases. Indeed, such a pattern can already be gleaned from the data. Figure

Figure 10: **Cumulative Number of Clinical Trials vs. Percentage of Population**



Notes: Data from ClinicalTrials.gov, accessed on 1/18/2024. We only consider the subset of lung cancer clinical trials that are interventional and not retrospective, observational, or real-world data (RWD). For consistency across alterations, we define "discovery" of an alteration the year of its first clinical trial. The percentage of lung cancer patients affected by each alterations on the horizontal axis is from Figure 1.

10 shows the relation between the cumulative number of clinical trials over a 10-year period since discovery of each alteration (EGFR, ALK, etc.) and the percentage of lung cancer population affected by each alteration. Even with such few observations, there is a nearly linear relation between the frequency of each alteration in the population and the number of clinical trials: Over the 10-year period after EGFR discovery, which affect about 14.5% of the lung cancer population, 103 clinical trials have been carried out. In contrast, during the 10-year period since the discovery of e.g. ALK, which affects about 5% of the population, only 29 clinical trials have been carried out. In an even sharper contrast, during the 10-year period since the discovery of the NTKR alteration, which affects 0.5% of the population, only 3 clinical trials have been carried out. Of course, there are many reasons why there have not been additional clinical trials for these rarer alterations. Our model though suggests that the current regulatory infrastructure is not well-suited to promote research to identify new drugs for ever smaller fractions of the population affected by individual genetic alterations.

5. Existing Solutions

Bayesian Adaptive Clinical Trials The costs related to clinical trials, especially for rare cancers or rare diseases, has spurred the medical community to experiment with different strategies. Bayesian adaptive clinical trials have been promoted as a methodology to overcome some of the issues discussed in previous sections. In a Bayesian adaptive clinical trial, the researcher starts with a prior distribution about the benefit of a new treatment (p_T in our model) versus the old treatment (p_C). Patients are recruited sequentially, and as new patients join the pool, the probability of the event that $p_T > p_C$ is updated. If the updated probabilities place sufficient weight on the event that either $p_T \ll p_C$ or $p_T \gg p_C$, then the trial is terminated and declared a failure or a success. Thus, the number of enrolled patients M may end up being smaller than in a regular RCT if the new treatments is in fact significantly superior or inferior than the old treatment. Conversely, if $p_T > p_C$ or $p_T < p_C$ only slightly, it may take longer to conclude the trial.

Bayesian adaptive clinical trials will only help in our model insofar as they decrease the total fixed R&D costs (e.g. through parallel testing), and the total number of patients M needed per alteration. Still, Table 3 suggests that even for very low costs of $R\&D$, as the number of alterations n increases, the break-even price becomes so high that it is unlikely that Bayesian adaptive clinical trials will make the economics sustainable for n large.

"Real-World" Data The medical profession has also been experimenting with a methodology that goes under the name Real World Data (RWD). This essentially entails using results from past clinical trials, as well as observational data on patient treatments, to compare new drugs to existing ones. In essence, given the small number of patients that can be enrolled in studies, one can run "single-arm" trials without control group and compare the results with those from previous trials (see Table 1 for examples of treatments that were approved with Single-Arm (SA) trials). While this strategy clearly reduces the number of patients needed, it does not solve the problem as n increases to a large number.

6. A New Paradigm: Approving the Drug Discovery Process

The existing solutions discussed above are driven by the same principle, namely that regulatory agencies approve each individual new drug. Our results show that this approach breaks down when scientific achievements make it increasingly more likely that medicine will approximate personalized treatment.

Recent advances in biomedicine (specifically, genome sequencing), the availability of relatively inexpensive computational power, and developments in artificial intelligence, make it possible to consider a novel alternative—a personalized drug discovery process. We first describe this process, and then discuss an approval regime for the drug development process as a whole, rather than for the individual drugs it identifies.

6.1. The Personalized Drug Discovery Process

The new drug discovery process would comprise the following steps:

A. Identification of driver alterations. For every tumor, there are thousands of alterations, some of which are critical to tumor growth. Other alterations (called "passengers") have no or modest impact on carcinogenesis. Each altered gene can generate an altered protein that can have a different function than in normal tissues, thus driving the carcinogenic process. Through the sequencing of tumor DNA and RNA, it is now possible to identify the majority of these critical drivers.

B. Rebuilding the altered protein. Methodologies have recently been developed that are able to rebuild the altered protein, a critical step in the personalized drug discovery process. Examples are Google DeepMind machine learning methods such as AlphaFold 3 or image-based single-cell functional tests. Evidence shows that such methods are able to rebuild the altered proteins with high level of precision. This process, which can now be done quickly in silico, used to take years in the past through experiments at the bench in wet laboratories.

C. Drug matching This step exploits super-computers to screen billions of molecules and identify those (new or existing) that can "block" the altered protein from signalling tumor growth, cancer cell survival and metastatization. In essence, the previous step determines the structure of the altered protein responsible for aberrant signalling. The drug-matching step instead involves the search of compounds that are in fact able to stop the signalling of the altered protein. This would still likely produce a large number of potential matches, but not all may be active enough or viable safe solutions for actual use.

D. Drug discovery Out of the many compounds created or selected in the previous step, the next task is to identify the ones that can be absorbed and tolerated by humans. The methodology here is to use the existing (and ever increasing) knowledge of the molecular structure of existing safe drugs, together with machine-learning methodologies, to find the compounds that are indeed safe for the organism. It is also possible that the an effective compound with the proper molecular structure may be found within the universe of existing drugs, and thus ensure its safe use.

E. Drug production Once the new compound has been identified, the new personalized drug must be physically produced. New automated machines, aided by 3D printing technologies, are considered to be viable technologies to produce ad-hoc drugs for individuals.

Teams around the world have already been implementing some of these steps. Ren et al. (2023) demonstrated for the first time that, by using Google Deepmind AlphaFold, steps (A) to (D) can be carried out to identify a new inhibitor within one month. They identified a novel hit molecule against a novel target for the treatment of hepatocellular carcinoma without any experimental structure. PandaOmics, a cloud-based software platform that applies AI and bioinformatics techniques, was employed to identify the protein of interest responsible for carcinogenesis (CDK20). Then, the authors leveraged Chemistry42, a comprehensive small-molecule drug discovery platform, to generate compounds that inhibit the activity of the protein, based on the structure predicted by AlphaFold. The selected molecules were synthesized and tested in biological assays. Through this approach, the authors identified a small molecule hit compound for the CDK20 alteration within 30 days from target selection, and after synthesizing 7 compounds only. A second round of AI-powered compound generation yielded a more potent hit molecule, ISM042-2-048. This new compound showed good CDK20 inhibitory activity and a selective anti-proliferation property.

Relatedly, Abramson et al. (2024) show that AlphaFold 3 is able to provide accurate structure prediction of biomolecular interactions. Finally, Kornauth et al. (2022) conducted a feasibility study and prospectively tested an image-based, single-cell functional approach to guide treatments in 143 patients with advanced aggressive hematologic cancers. 56 patients (39%) were treated according to functional results. In 54% of patients, there was a demonstrated clinical benefit over traditional therapies.

Beyond cancer treatment, Zhang et al. (2025) provides a general overview of advances in the drug development process that have been driven by AI. In particular, the authors highlight how AI is accelerating all stages of drug development, from target identification (our step A) to drug discovery (steps B through D). Moreover, recently the FDA issued a new draft guidance and call for comments on AI-driven drug development (see https://www.fda.gov/media/184830/download) which further validates the new methodologies. To sum up, from a science perspective, the personalized drug development process we analyze aligns with broad trends in biomedicine; our objective in this section is to highlight its economic implications.

As regards step E, Gao et al. (2021) and Wang et al. (2023) discuss 3D printing technologies for personalized drug production. In 2016, for instance, the FDA approved the first drug produced by 3D printing (SPRITAM). While no targeted therapy cancer drug has yet been successfully produced by 3D printing, it is conceivable that progress in these technologies

will make it possible to produce such drugs in the future.

6.2. Economic Viability

Would such personalized drug-discovery process be viable from an economic standpoint? Consider a new patient arriving at a hospital with a newly diagnosed tumor. This is when the personalized drug discovery process would start. We now analyze its economics.

Let $C_{PDDP}(N)$ be the cost per patient of going through steps A through E of the PDDP process. The cost per patient depends on the total number N of patients, as it may include e.g. fixed costs. With probability Q, the discovery process is successful—that is, a personalized treatment is found. Let $P^{NewDrug}$ be the price charged to the patient for this personalized treatment, and $c^{NewDrug}$ be the production cost (step E). Denote the survival probability by s^{new}. Unless $s^{new} = 1$, there is still the possibility that another, better personalized treatment may be found over time for the same patient; we denote its annual probability π^{new}. Now consider the case in which the pharma company fails to identify a personalized treatment; this occurs with probability $1-Q$. In this case, the patient is treated using the best existing alternative, at a price for the patient of $P^{OldDrug}$, and a cost to the company of $c^{OldDrug}$.

The present value for a (monopolistic) pharma company of carrying out the discovery process as just described for the entire population of N patients per year is thus:

$$\begin{aligned} PV_{PDDP} &= Q \times A^{New} \times (P^{NewDrug} - c^{NewDrug}) \times N - C_{PDDP}(N) \times N \qquad (16) \\ &\quad + (1-Q) \times A^{Old} \times (P^{Old} - c^{Old}) \times N. \qquad (17) \end{aligned}$$

Here,

$$A^{New} = \frac{1-\pi^{new}}{R+\pi^{new}} \left(\frac{1+R}{1+R-(1-\pi^{new}) s^{new} s^{other}} \right) \qquad (18)$$

is the discount factor from the revenues generated by the sale of the personalized drug to the patient.[20] This is determined by the survival probability s^{new}, the probability π^{new} of a new, better treatment being discovered, and the survival probability s^{other} related to death from causes other than cancer.

The last line in the expression for PV_{PDDP}, i.e., Eq. (17), accounts for the fall-back plan in case the drug discovery process does not deliver a new drug. In this case, the patient is treated using one of the existing (but still targeted) drugs. In the expression, A^{Old} is the same as A^{New} except that the survival rate s and probability of competitor π depend on which "Old drug" is used. Note that this present value takes into account the present

[20] This discount factor is equal to the one in Eq. (4) except that $\tau = 0$ as there is no RCT in this case.

value of profits across all patients, but it does not involve the approval from the FDA, as the process itself is assumed to have been approved already (see Section 6.5.).

6.3. The Costs of Personalized Drug Discovery Process

An important element in this calculation is whether the cost component $C_{PDDP}(N)$ is sufficiently small to generate a positive present value for the Pharma industry as a whole. These costs should also include the "rental cost" of supercomputers[21] to perform these calculations. Indeed, the main cost of personalized drug development may likely be the software/hardware component, given that most of the steps involve a computer-based discovery process, except for the initial steps in genome sequencing. It is plausible to expect increasing return to scale initially: the same supercomputer can be used to find compounds for all patients, so costs are amortized over large numbers of patients. This will eventually lower the cost per patient. Regardless, the critical point is that costs are independent of the number n of alterations.

6.3.1. Argonne Aurora Supercomputer

We can obtain rough estimates of $C_{PDD}(N)$ by calculating the costs of purchasing a supercomputer and running year by year. Of course, this is a speculative exercise, but it is useful to understand the order of magnitude of the costs involved. In addition, our estimates are based on the current state of technology, and it is likely that costs will decrease over time.

As a case in point, consider the the Aurora supercomputer at the Argonne National Laboratory. The cost of this computer is estimated at $500 million and it is predicted to consumes 60 Megawatts once at peak performance.[22] According to Vasan et al. (2023), with only 256 nodes (out of 9000) of Aurora researchers were able to screen 2 billion molecules per hour. Assuming that number of molecules screened increases linearly with number of nodes deployed, researchers expect to be able to "screen up to a trillion of compounds per hour if using all of Aurora's computer resources" (Heinonen, 2024).

Assuming that the screening of 22 billion compounds is sufficient (on average) to find the compound for one patient, then we can expect to be able to screen 306,600 patients per year.[23] Conservatively, we can thus assume that an Aurora computer can screen the entire population of new patients per year with targetable cancer (110,000 patients per year in the US). Assuming conservatively a depreciation rate of 100% over 4 years, this implies

[21] A supercomputer is a high-performance computing machine designed to perform complex calculations at extremely high speeds. It typically consists of thousands of interconnected processors working in parallel to solve large-scale computational problems that are beyond the capabilities of standard computers.

[22] https://en.wikipedia.org/wiki/Aurora_(supercomputer) accessed on 10/12/2024.

[23] Aurora has 9000 nodes. Therefore, assuming linear scalability, 9000/256 = 35 patients/hour. Thus, 35/hour ×24 hours × 365 = 306,600 patients/year.

that Aurora could screen 440,000 patients over its lifetime, and hence, the cost per patient would be \$1136/patient. In addition, Aurora consumes 60 MW at peak. Assuming the cost of $\$60 MWh$, we obtain that the electricity cost of running Aurora is \$31.5 million/year.[24]. Dividing by 110,000 patients we obtain an electricity cost of \$286 per patient.

All in all, we can estimate a cost of \$1422/patient by dedicating an Aurora supercomputer to personalized drug discovery process.

6.3.2. AlphaFold

Google DeepMind AlphaFold has been successful in predicting molecular structures that are fundamental for the implementation of the Personalized Drug Discovery process. A difficulty in assessing the true costs of using this technology is that, at present (February 2025), researchers can access AlphaFold technology at essentially no cost for research purposes. We would expect this to change with the implementation of a for-profit Personalized Drug Discovery Process at scale.

However, a recent commentary paper by Callaway (2024) mentions that an open-source replacement for AlphaFold3 is under development, and that training such an open-source model is expected to cost about \$1 million. Raouraoua et al. (2024) introduced "Massive-Fold," which improved the computing speed of AlphaFold by "reducing the computing time from months to hours." Pairing such models with a supercomputers like Aurora should provide cost-effective strategies to support the Personalized Drug Discovery Process.

6.4. Pharma Companies' Incentives

Consider the pharma industry as a whole. The present value of offering the personalized drug discovery process to patients is given by Eq. (16). Alternatively, the pharma industry may continue to offer the existing drugs, which yield a present value of $PV_{existing} = A^{Old} \times (P^{Old} - c^{Old}) \times N$; as above, we assume that no new clinical trials and/or discovery costs are necessary for existing treatments. We can calculate the break-even price to ensure that the Pharma industry is indifferent between moving to the Personalized Drug Discovery Process and the current therapies. That is, we seek the price $P^{NewDrug}$ such that $PV_{PDDP} = PV_{existing}$.

Two additional parameters are hard to gauge at this time: the probability Q that steps (A) through (E) will deliver a drug, and the associated survival probability s^{new} under the new process. The survival probability affects the discount factor A^{New}, as the longer the patient lives, the longer the pharma company can deliver the drug to the patient. Moreover, it affects the probability that a better drug is discovered: If the survival probability from

[24]This is computed as \$60/MWh $\times$ 24 $\times$ 365 $\times$ 60MW = \$31,5 million.

the personalized drug is $s^{new} = 100\%$, then clearly the probability π^{new} of finding a better drug drops to zero. The incentives are thus aligned for the pharma company to find the best possible drug for that particular patient.

To circumvent this difficulty, we calculate the break-even margins for various combinations of Q and s^{new}, with the additional assumption that a higher s^{new} also implies a lower probability π^{new} of finding a better drug in the future. Specifically, we assume that $\pi^{new} = 30\%$ if $s^{new} = 50\%$ and $\pi^{new} = 0$ if $s^{new} = 100\%$, and we interpolate for intermediate values. The break-even margin is

$$(P^{NewDrug} - c^{NewDrug}) = \frac{C_{PDDP}(N) + Q \times A^{Old} \times (P^{Old} - c^{Old})}{Q \times A^{New}} \tag{19}$$

As in the simulations in Section 4., we assume a margin per patient for treatment using existing drugs of \$4000/week, for a total of $(P^{Old} - c^{Old}) = \$208,000$ per year, per patient. We also set the annual survival probability for the existing treatment to s^{old}=50%, halfway between s_C and s_T in Table 2. Intuitively, the new PDDP will "compete" with existing targeted therapies. In order for the new PDDP to compete with this profitable business, the new drug must cover revenues lost from replacing the old treatment (whose margin we hold fixed) with the new, as well as the additional costs to run the PDDP process. The values of the break-even margin are in Table 4. A higher probability of success Q, and especially a higher survival probability s^{new}, lead to a lower annual break-even margin per patient. Intuitively, the higher the survival probability for the new personalized drug, the longer the pharma company can receive revenues from it, and the less likely it is that a better drug can be found. Both of these effects increase the discount factor A^{New}, and hence substantially reduce the break-even margin. Similarly, a higher probability Q of finding a personalized drug also decreases the break-even margin, as it makes it more likely the pharma company will be able to sell the new drug for a longer period.

The values in Table 4 should be compared with the current cost of targeted therapies, which we assume to be $\$208,000$ per year. As the survival probability increases, the potential savings per patient are considerable. Even for $s^{new} = 60\%$, a realistic number given that we set $s^{old} = 50\%$ and that $s_T = 64\%$ for the new treatment in Table 4, the breakeven margin is just over \$145,000 even for a probability of success of only $Q = 10\%$. When s^{new} is 70%, the break-even margin to induce pharma companies to switch to PDDP is just over $\$90,000$, less than half the assumed base cost of targeted therapies of $\$208,000$.

6.5. Approval of the Personalized Drug Discovery Process

The personalized drug discovery process described above is intended to provide new drugs that are effective and safe for humans. This needs to be established in a randomized control

Table 4: Break-even Margin $P^{NewDrug} - c^{NewDrug}$

Probability of new	survival probability s^{new}				
personalized treatment Q	60%	70%	80%	90%	100%
10%	\$145,848	\$93,034	\$53,338	\$25,294	\$7,784
20%	\$143,885	\$91,782	\$52,621	\$24,953	\$7,679
30%	\$143,231	\$91,365	\$52,382	\$24,840	\$7,644
40%	\$142,904	\$91,156	\$52,262	\$24,783	\$7,626
50%	\$142,708	\$91,031	\$52,190	\$24,749	\$7,616
60%	\$142,577	\$90,948	\$52,142	\$24,726	\$7,609
70%	\$142,484	\$90,888	\$52,108	\$24,710	\$7,604
80%	\$142,414	\$90,843	\$52,083	\$24,698	\$7,600
90%	\$142,359	\$90,808	\$52,063	\$24,689	\$7,597
10%	\$142,316	\$90,781	\$52,047	\$24,681	\$7,595

trial itself. The randomized clinical trial in this case would be set up in the same format as in Section 3.. In particular, consider a number of new patients M, which we divide in two groups: The test T group and the control C group. The C group will receive the best existing drug from the start. The T group undergoes the steps A through E in Section 6.1. for the personalized drug discovery process (including, if necessary, the existing treatment as a fall-back option). The patients in the two groups are then followed for τ periods. Note that in this case we do not differentiate across alterations. All patients with lung cancer are considered as just one group. All other steps are otherwise the same, and after the observation period of τ periods the usual calculations are performed.

7. Discussion and Conclusions

In this paper, we focused on lung cancer as an example of the economic infeasibility of personalized medicine under the current approval regime for drugs. We chose lung cancer both because it is a relatively common diseases, and because considerable progress has been made on discovering new biomarkers and developing targeted therapies. Our main finding is that, despite these recent scientific and therapeutical breakthroughs, as the number of alterations increase, it will become either uneconomical or prohibitively costly for society to continue developing new drugs under the current approval regime.

Our theoretical results apply well beyond lung cancer treatments. Indeed, considerable progress in targeted therapies has been made for other types of cancer as well, including breast, colon, ovarian cancers, as well as hematological malignancies. Similar progress in personalized medicine has also been made in the last decade in other areas of medicine,

including hematology, cardiology, neurology, and more. Recent research has pushed personalized medicine into the fight against antibiotic resistance, for instance.

The general point, however, is that personalized medicine is akin to turning all diseases into rare diseases, which discourages discovery. Yet, new technologies based on machine learning and artificial intelligence now allow for personalized drugs to be developed in silico. We show that regulatory agencies' approval of the *process* for drug discovery, as opposed to individual therapies, can avoid the difficulties and the large costs undergoing the different phases of the traditional drug discovery process (Phase I, II, and III) thereby providing the proper incentives for researchers and pharma companies to develop new personalized treatments. Indeed, we show that current technologies would make it economical for pharma companies to switch to a Personalized Drug Development Process (PDDP), with substantial gains for society. Moreover, assuming that technological advancements will continue, it is also likely that the costs of PDDP per person will further decline in the future.

Finally, we have assumed that for a given disease (e.g. lung cancer) all genomic alterations have identical incidence in the population, which decreases as research finds additional alterations. In reality, some alterations are rarer than others. Our results thus suggest that the current approval process will eventually hamper the discovery of targeted therapies for rare alterations, generating an asymmetry between patients with the same disease: Eventually, those with the rarest mutations will not have targeted therapies available, while others with more common alterations will.

References

[1] Josh Abramson et al. "Accurate structure prediction of biomolecular interactions with AlphaFold 3". In: *Nature* (2024), pp. 1–3.

[2] Eric Budish, Benjamin N Roin, and Heidi Williams. "Do firms underinvest in long-term research? Evidence from cancer clinical trials". In: *American Economic Review* 105.7 (2015), pp. 2044–2085.

[3] Ewen Callaway. "Who will make AlphaFold3 open source? Scientists race to crack AI model". In: *Nature* 630.8015 (2024), pp. 14–15.

[4] D Ross Camidge et al. "Brigatinib versus crizotinib in ALK inhibitor–naive advanced ALK-positive NSCLC: final results of phase 3 ALTA-1L trial". In: *Journal of Thoracic Oncology* 16.12 (2021), pp. 2091–2108.

[5] Daniel Carpenter and Michael M Ting. "Regulatory errors with endogenous agendas". In: *American Journal of Political Science* 51.4 (2007), pp. 835–852.

[6] Joseph A DiMasi, Henry G Grabowski, and Ronald W Hansen. "Innovation in the pharmaceutical industry: new estimates of R&D costs". In: *Journal of health economics* 47 (2016), pp. 20–33.

[7] Alexander P Frankel, Joshua L Krieger, Danielle Li, and Dimitris Papanikolaou. *Evaluation and Learning in R&D Investment*. Tech. rep. University of Chicago, 2024.

[8] Ge Gao et al. "3D printing of pharmaceutical application: drug screening and drug delivery". In: *Pharmaceutics* 13.9 (2021), p. 1373.

[9] Daniel A Goldstein et al. "A global comparison of the cost of patented cancer drugs in relation to global differences in wealth". In: *Oncotarget* 8.42 (2017), p. 71548.

[10] Yingni Guo. "Dynamic delegation of experimentation". In: *American Economic Review* 106.8 (2016), pp. 1969–2008.

[11] Nils Heinonen. *Using molecular deep learning models on Aurora to accelerate drug discovery*. 2024.

[12] Raymond H Henderson et al. "Delivering the precision oncology paradigm: reduced R&D costs and greater return on investment through a companion diagnostic informed precision oncology medicines approach". In: *Journal of Pharmaceutical Policy and Practice* 16.1 (2023), p. 84.

[13] Lizza EL Hendriks et al. "Non-small-cell lung cancer". In: *Nature Reviews Disease Primers* 10.1 (2024), p. 71.

[14] Emeric Henry, Marco Loseto, and Marco Ottaviani. "Regulation with experimentation: Ex ante approval, ex post withdrawal, and liability". In: *Management Science* 68.7 (2022), pp. 5330–5347.

[15] Emeric Henry and Marco Ottaviani. "Research and the approval process: The organization of persuasion". In: *American Economic Review* 109.3 (2019), pp. 911–955.

[16] Takakazu Honryo. "Dynamic persuasion". In: *Journal of Economic Theory* 178 (2018), pp. 36–58.

[17] International Human Genome Sequencing Consortium. "Initial sequencing and analysis of the human genome". In: *nature* 409.6822 (2001), pp. 860–921.

[18] IQVIA. *Global Oncology Trends 2024: Outlook to 2028*. Tech. rep. IQVIA Institute, 2024. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/global-oncology-trends-2024.

[19] Emir Kamenica and Matthew Gentzkow. "Bayesian persuasion". In: *American Economic Review* 101.6 (2011), pp. 2590–2615.

[20] Christoph Kornauth et al. "Functional precision medicine provides clinical benefit in advanced aggressive hematologic cancers and identifies exceptional responders". In: *Cancer discovery* 12.2 (2022), pp. 372–387.

[21] Natasha B Leighl et al. "An arm and a leg: the rising cost of cancer drugs and impact on access." In: *American Society of Clinical Oncology Educational book. American Society of Clinical Oncology. Annual Meeting*. Vol. 41. 2021, pp. 1–12.

[22] Andrew W Lo and Shomesh E Chaudhuri. *Healthcare Finance: Modern Financial Analysis for Accelerating Biomedical Innovation*. Princeton University Press, 2022.

[23] Andrew McClellan. "Experimentation and approval mechanisms". In: *Econometrica* 90.5 (2022), pp. 2215–2247.

[24] Myriam Mirza et al. "Assessing the Cost-Effectiveness of Next-Generation Sequencing as a Biomarker Testing Approach in Oncology and Policy Implications: A Literature Review". In: *Value in Health* (2024).

[25] Tony Mok et al. "Updated overall survival and final progression-free survival data for patients with treatment-naive advanced ALK-positive non-small-cell lung cancer in the ALEX study". In: *Annals of oncology* 31.8 (2020), pp. 1056–1064.

[26] Thomas J Moore et al. "Variation in the estimated costs of pivotal clinical benefit trials supporting the US approval of new therapeutic agents, 2015–2017: a cross-sectional study". In: *BMJ open* 10.6 (2020), e038863.

[27] Nessim Raouraoua et al. "MassiveFold: unveiling AlphaFold's hidden potential with optimized and parallelized massive sampling". In: *Nature Computational Science* (2024), pp. 1–5.

[28] Feng Ren et al. "AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor". In: *Chemical science* 14.6 (2023), pp. 1443–1452.

[29] Jacqulyne P Robichaux et al. "Structure-based classification predicts drug response in EGFR-mutant NSCLC". In: *Nature* 597.7878 (2021), pp. 732–737.

[30] Robert Roskoski Jr. "Properties of FDA-approved small molecule protein kinase inhibitors: A 2024 update". In: *Pharmacological research* (2024), p. 107059.

[31] Aylin Sertkaya et al. "Costs of drug development and research and development intensity in the US, 2000-2018". In: *JAMA Network Open* 7.6 (2024), e2415445–e2415445.

[32] Karen E Skinner et al. "Healthcare costs in patients with advanced non-small cell lung cancer and disease progression during targeted therapy: a real-world observational study". In: *Journal of medical economics* 21.2 (2018), pp. 192–200.

[33] Benjamin J Solomon et al. "Lorlatinib Versus Crizotinib in Patients With Advanced ALK-Positive Non–Small Cell Lung Cancer: 5-Year Outcomes From the Phase III CROWN Study". In: *Journal of Clinical Oncology* (2024), JCO–24.

[34] Tarek Taha et al. "Treatment of rare mutations in patients with lung cancer". In: *Biomedicines* 9.5 (2021), p. 534.

[35] Archit Vasan et al. "Scalable Lead Prediction with Transformers using HPC resources". In: *Proceedings of the SC'23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*. 2023, pp. 123–123.

[36] Gregory A Vidal et al. *Practice-and provider-level inequities in next-generation sequencing (NGS) testing by race/ethnicity for patients (pts) with advanced non-small cell lung cancer (aNSCLC) treated in the community setting.* 2023.

[37] Shanshan Wang et al. "A review of 3D printing technology in pharmaceutics: Technology and applications, now and future". In: *Pharmaceutics* 15.2 (2023), p. 416.

[38] Chi Heem Wong, Kien Wei Siah, and Andrew W Lo. "Estimation of clinical trial success rates and related parameters". In: *Biostatistics* 20.2 (2019), pp. 273–286.

[39] Kang Zhang et al. "Artificial intelligence in drug development". In: *Nature Medicine* (2025), pp. 1–15.

Supplemental Appendix to Accompany the Paper

On the Economics Infeasibility of Personalized Medicine, and a Solution

Marina Chiara Garassino
University of Chicago Comprehensive Cancer Center

Kunle Odunsi
University of Chicago Comprehensive Cancer Center

Marciano Siniscalchi
Northwestern University

Pietro Veronesi
University of Chicago, NBER, and CEPR

A Appendix: Monte Carlo Simulations

In this appendix we describe the procedure to obtain the results in the calibration section. We perform all calculations using the statistics for small-sample distributions through Monte Carlo simulations. In particular, for a given sample M of patients, with $M/2$ in the treatment and control group, and with x_T and x_C as the number of successes for the treatment and control group, respectively, the distribution under the null hypothesis $p_T = p_C$ can be written as the hypergeometric function:

$$Prob\left(X_T = x_T | X_T + X_C = t, H_0\right) = \frac{\left(\begin{array}{c} M/2 \\ x_T \end{array}\right)\left(\begin{array}{c} M/2 \\ t - x_T \end{array}\right)}{\left(\begin{array}{c} M \\ t \end{array}\right)}. \tag{20}$$

The Exact Fisher Test rejects the null hypothesis $p_T = p_C$ whenever the number of successes in the treatment group x_T is larger than a cutoff point $c(t)$, i.e. $x_T > c(t)$, which can be obtained by inverting the hypergeometric function in equation (20) at the desired level of confidence $1 - \alpha$.

For given n, we then proceed as follows:

- For each M running from $M = 2$ to $M = N$ (only even numbers) we simulate $Sim = 100,000$ samples of M uniformly distributed random variables. We also simulate $Sim = 100,000$ samples of random variable $\tilde{p}_T$ from the truncated normal distribution described above. Denote the realization of $\tilde{p}_T$ in sample sim as p_T^{sim}.

- We divide each sample in two groups of size $M/2$, and simulate the number of successes (1's) and fails (0's) for each subgroup according to the probabilities p_T^{sim} and p_C, where the former is itself simulated from the truncated normal distribution. Denote by x_T^{sim} and x_C^{sim} the number of successes for each group, respectively, in simulation sim.

- Determine $t^{sim} = x_T^{sim} + x_C^{sim}$ and compute the 95% cutoff $c(t^{sim})$ by inverting the hypergeometric cumulative distribution in equation (20).[25]

- If $x_T^{sim} > c(t^{sim})$ we reject the null hypothesis that $p_T = p_C$, and thus assume that FDA will accept the new treatment.

- The ex-ante probability of rejection can be computed as

$$Pr(Approval|M,n) = E\left[1_{x_T>c(t)}\right] = \frac{1}{Sim}\sum_{sim=1}^{Sim} 1_{x_T^{sim}>c(t^{sim})}$$

  where $1_{x_T>c}$ denotes the indicator function for $x_T > c$.

- Given the $Pr(Approval|M,n)$ for each M, we can now move to compute the $Pr(Approval|K,n)$ by using expression

$$\Pr\left(Approval|K,n\right) = 1 - \sum_{M=2}^{K} Pr(M|K)Pr(Approval|M,n) \tag{21}$$

  where recall that $Pr(M|K,n)$ is given by expression (5).

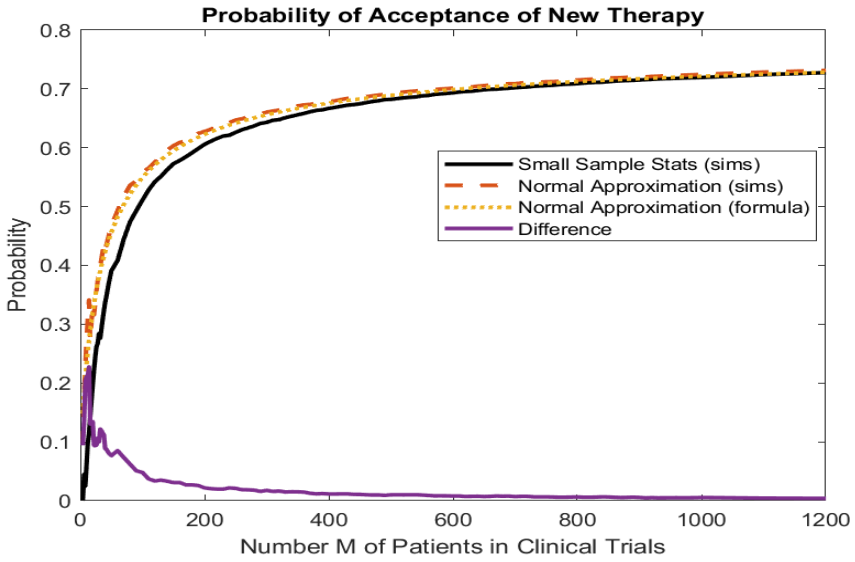
- We then repeat the whole procedure above for $n = 2$ to $n = 10000$.

To verify the simulation procedure and that the number of simulations $Sim = 100,000$ are sufficient, we perform the same calculations as above using the normal distribution approximation; the only difference is that the rejection of the null hypothesis H_0 is computed using the cutoff rule (1) for each simulated sample.

Figure 11 shows that indeed, the exact Fisher test is more likely to reject the new treatment for small number of patients compared to the normal distribution. However, the two distributions converge as the number M of patients in clinical trial increases. The figure also shows that the number of simulations are sufficient to provide accurate statistics, as the case with simulated normal random variables overlays the one from the theoretical formula.

[25] The inverse of the hypergeometric cumulative distribution is available in standard statistical packages. We use Matlab with $c(t) = hygeinv(.95, M, M/2, t)$.

Figure 11: **Probability of Acceptance of New Treatment**



A1. Standardized Marginal Improvement and Success Uncertainty

As noted at the end of Section 3.3., the fact that new targeted treatments are compared to the best available treatment at the time of the trial implies that the standardized marginal improvement κ may either decrease or increase as the number of alterations n increases. We now illustrate this point using recent developments in treatments for ALK alterations as a concrete example.

Column 3 in Table 5 shows the progression of treatment of patients with this alteration (see T. Mok et al., 2020, D Ross Camidge et al., 2021 and Solomon et al., 2024). As it can be seen, since 2013, the 3-year progression free survival (PFS) has increased from 5% with chemotherapy, to 20% with the first targeted therapy in 2013 (Crizotinib), and all the way to 65% with Lorlatinib in 2021 (we inverted Brigatinib with Alectinib in the chronological order to rank the treatment by progression-free survival rate in column 3.) When the benchmark was chemotherapy with PFS = 5%, the standardized marginal improvement κ could be rather large even for treatments with small improvements in PFS. Taking $p_C = PFS$ as our measure of success, Columns 4 and 5 of Table 5 show the standardized marginal improvement κ for two assumptions about the marginal improvements of each new therapy: $\Delta p = p_T - p_C =$ 10% and 15%. When the benchmark is chemotherapy, κ is 0.24 and 0.33, respectively, for the two scenarios. Consider now the benchmark as being Crizotinib, which yields a $PFS = p_C = 20.2\%$. Now, new treatments that improve over the new benchmark by the same margins (i.e. 10% and 15%) would have κ equal 0.16, and 0.24, respectively, which are far lower than they would be if chemotherapy was the benchmark. Indeed, Alectinib became

Table 5: ALK Alteration and Approval Probabilities

			Standardized marginal improvement κ for two $\Delta p = p_T - p_C$		Probability of approval with sample $M = 400$ and			
					no uncertainty		uncertainty on p_T	
	Date	3-year PFS (p_C)	$\Delta p = 10\%$	$\Delta p = 15\%$	10%	15%	10%	15%
Chemo	–	5.0 %	0.24	0.33	92%	100%	72%	81%
Crizotinib	2013	20.2%	0.16	0.24	64%	93%	55%	65%
Brigatinib	2018	43.0%	0.14	0.21	52%	86%	51%	61%
Alectinib	2015	50.6%	0.14	0.22	53%	87%	51%	60%
Lorlatinib	2021	65.0%	0.16	0.24	59%	93%	53%	57%
Hypothetical		75.0%	0.18	0.28	71%	98%	52%	51%

the benchmark in 2018; under this benchmark, the κ's would decrease further for the same marginal improvement. A lower κ implies a lower probability of acceptance, as shown in equation (2), and thus a higher required number of patients for the randomized clinical trial.

As the benchmark p_C increases, a second issue comes to play: any uncertainty around the success of the new treatment p_T further decreases the probability of acceptance for given sample size, and thus the incentives to develop new drugs. To see this in a simple example, consider first the probability of acceptance using a sample size of $M = 400$ patients. With no uncertainty about p_T and using formula (2), the probabilities of acceptance are as in columns 6 and 7 in Table 5. Notice that the probability of acceptance is monotonic in κ. In particular, for given M, the probability of acceptance follows a U-shape, with the minimum probability of acceptance being for intermediate values of p_C.

Assume now that p_T is a truncated normal distribution with $\tilde{p}_T \sim TrN(p_T, \sigma^2_{p_T}, 0, 1)$. We set $\sigma_{P_T} = (p_T - p_C)/0.55$: that is, $Pr(p_T < p_C) = 30\%$ under the normal distribution. Columns 8 and 9 of Table 5 now show a nearly monotonically declining pattern of the probability of acceptance as the benchmark p_C increases. As p_C increases and $p_T = p_C + \Delta p$ increases, we are shifting some mass to the left of distribution. In essence, it becomes increasingly less likely to beat the benchmark. While this is just an example that depends on our parametric assumptions, the general point is that there need not be a tight relation between κ and the probability of acceptance when there is uncertainty about p_T. In the limit, if the baseline PFS is $p_C = 90\%$, it becomes very unlikely to find a successful competing drug with $p_T > 90\%$.

B Proofs

Throughout this section, let

$$g(M) \equiv \begin{cases} \Phi\left(z(\alpha) - \kappa\sqrt{\frac{M}{2}}\right) & M \geq 0 \\ \Phi(z(\alpha)) & M < 0. \end{cases} \tag{22}$$

B1. Proof of Proposition 1

For the remainder of this appendix, denote by E[·], Var[·], and Cov[·, ·] the expectations, variances and covariances relative to the approximating density $\phi_{Kq,Kq(1-q)}$ with mean Kq and variance $Kq(1-q)$. Thus, in particular,

$$PV(K,n) = (1 - \mathrm{E}[g(M)]) \cdot C - K \cdot c_{RCT} \tag{23}$$

where

$$C = A(P_{TT} - c)N. \tag{24}$$

To prove Proposition 1, we show that $\partial/\partial K E[g(M)] < 0$, which implies the claim. By the Normal Approximation assumption, we can write

$$E\left[g(M)\right] = \int \frac{1}{\sqrt{2\pi Kq(1-q)}} e^{-\frac{1}{2Kq(1-q)}(M-Kq)^2} g(M)\, dM.$$

Differentiating wrto K yields

$$\begin{aligned}
&\int \frac{-\frac{1}{2}K^{-\frac{3}{2}}}{\sqrt{2\pi q(1-q)}} e^{-\frac{1}{2Kq(1-q)}(M-Kq)^2} g(M)\, dM \\
&+ \int \frac{1}{\sqrt{2\pi Kq(1-q)}} e^{-\frac{1}{2Kq(1-q)}(M-Kq)^2} \left\{ -\frac{-K^{-2}}{2q(1-q)}(M-Kq)^2 - \frac{1}{2Kq(1-q)} 2(M-Kq)(-q) \right\} \\
&\quad \cdot g(M)\, dM \\
&= \int \left(-\frac{1}{2K}\right) \phi_{Kq,Kq(1-q)}(M) g(M)\, dM \qquad (25) \\
&+ \int \left(\frac{1}{2K^2 q(1-q)}(M-Kq)^2\right) \phi_{Kq,Kq(1-q)}(M) g(M)\, dM \qquad (26) \\
&+ \int \left(\frac{1}{K(1-q)}(M-Kq)\right) \phi_{Kq,Kq(1-q)}(M) g(M)\, dM. \qquad (27)
\end{aligned}$$

Adding up the terms in parentheses in Eqs. (25)–(27) yields

$$\left[-Kq(1-q) + (M-Kq)^2 + 2Kq(M-Kq)\right] \cdot \frac{1}{2K^2 q(1-q)}.$$

Moreover,

$$\begin{aligned} &- Kq(1-q) + (M - Kq)^2 + 2Kq(M - Kq) \\ &= - Kq(1-q) + M^2 - 2MKq + K^2q^2 + 2MKq - 2K^2q^2 \\ &= - Kq(1-q) + M^2 - K^2q^2 \\ &= - \text{Var}[M] + M^2 - \text{E}[M]^2 = M^2 - \text{E}[M^2]. \end{aligned} \tag{28}$$

Therefore,

$$\frac{\partial E[g(M)]}{\partial K} = \frac{\text{E}\left[(M^2 - \text{E}[M^2]) \cdot g(M)\right]}{2K^2q(1-q)} = \frac{\text{Cov}\left[M^2, g(M)\right]}{2K^2q(1-q)} < 0 \tag{29}$$

because $g(M)$ is decreasing in M. *Q.E.D.*

B2. Proof of Proposition 2

The FOC for the optimal choice of K given q in Eq. (23) is

$$\frac{\partial PV_{TT}(n)}{\partial K} = -\frac{\partial E[g(M)|K,q]}{\partial K} C - c_{RCT} = 0. \tag{30}$$

By the implicit function theorem,

$$\frac{\partial K}{\partial C} = -\frac{-\frac{\partial E[g(M)|K,q]}{\partial K}}{-\frac{\partial^2 E[g(M)|K,q]}{\partial K^2} \cdot C}.$$

By assumption, PV is concave in K, and the denominator of the above fraction equals $\partial^2 PV_{TT}(n)/\partial K^2$, so it is negative. By Proposition 2, the numerator is positive, so overall $\partial K/\partial C > 0$: that is, the optimal choice of K is increasing in C for fixed c_{RCT}. Since $C = A(P_{TT} - c)N$, it follows that K^* increases if $P_{TT} - c)N$, i.e., profits per period, increase, or if A is higher. The latter occurs if R τ, or π are lower, or if s is larger.

Furthermore,

$$\frac{\partial K}{\partial c_{RCT}} = -\frac{-1}{-\frac{\partial^2 E[g(M)|K,q]}{\partial K^2} \cdot C}$$

which is negative overall because, as argued above, the denominator is negative. Thus, K^* increases if c_{RCT} decreases.

Finally,

$$\frac{\partial K}{\partial \kappa} = -\frac{-\frac{\partial^2 E[g(M)|K,q]}{\partial K \partial \kappa}}{-\frac{\partial^2 E[g(M)|K,q]}{\partial K^2}} = \frac{\frac{\partial^2 E[g(M)|K,q]}{\partial K \partial \kappa}}{-\frac{\partial^2 E[g(M)|K,q]}{\partial K^2}};$$

as argued above, since $C > 0$, the denominator is negative. In the numerator, from Eq. (29),

$$\frac{\partial \text{E}[g(M)|K,q]}{\partial K} = \frac{\partial \text{E}\left[(M^2 - \text{E}[M^2]) \cdot g(M)\right]}{2K^2q(1-q)}$$

and so, differentiating wrto κ and noting that $g(M) = \Phi(z(\alpha))$ for $M < 0$ yields

$$\frac{\partial^2 E[g(M)|K,q]}{\partial K \partial \kappa} = \frac{1}{2K^2 q(1-q)} \cdot \frac{\partial \mathrm{E}\left[(M^2 - \mathrm{E}[M^2]) \cdot g(M)\right]}{\partial \kappa} =$$

$$= \frac{1}{2K^2 q(1-q)} \cdot \int_0^\infty (M^2 - \mathrm{E}[M^2]) \cdot \phi\left(z(\alpha) - \kappa\sqrt{\frac{M}{2}}\right)\left(-\sqrt{\frac{M}{2}}\right) \phi_{Kq,Kq(1-q)}(M)\, dM =$$

$$= -\frac{1}{2K^2 q(1-q)} \cdot \int_{-\infty}^\infty (M^2 - \mathrm{E}[M^2]) \cdot \phi\left(z(\alpha) - \kappa\sqrt{\frac{M^+}{2}}\right)\left(\sqrt{\frac{M^+}{2}}\right) \phi_{Kq,Kq(1-q)}(M)\, dM$$

where $M^+ = \max(M, 0)$. The integral in the above expression is equal to $\mathrm{Cov}[M^2, \phi(z(\alpha) - \kappa\sqrt{M^+/2})\sqrt{M^+/2}]$. If this covariance is negative, the numerator is positive, and so, since the denominator is negative, $\partial K/\partial \kappa < 0$ as claimed. *Q.E.D.*

B3. Proof of Proposition 3

Under Eq. (11), we replace the term $(1-q)$ in the approximating density of M with 1. Throughout the proof of Proposition 1, the term $(1-q)$ is treated as a constant; specifically, in the derivation beginning with Eq. (28), the term $Kq(1-q)$ is never multiplied out. Replacing $(1-q)$ with 1 thus does not affect the argument, and furthermore, Kq is indeed the variance of M under the Normal Approximation assumption 2. Hence, the argument in the proof of Proposition 1 also implies that

$$\frac{\partial}{\partial K} \int \frac{1}{\sqrt{2\pi Kq}} e^{-\frac{1}{2Kq}(M-Kq)^2} g(M)\, dM < 0.$$

In the above integral, K and q are symmetric, so a symmetric argument shows that the integral is also decreasing in q, i.e., increasing in n. Under the Normal Approximation 2, the probability of approval $p(K,n)$ is 1 minus the above integral, and thus it is decreasing in n, as claimed. *Q.E.D.*

B4. Proof of Proposition 4

Let $n > n'$ and denote by K and K' the optimal choices of sample sizes for n and n', respectively. By Proposition 3,

$$PV_{III}(K,n) = p(K,n) \cdot C - K \cdot c_{RCT} < p(K,n') \cdot C - K \cdot c_{RCT} = PV_{III}(K,n');$$

furthermore, since K' is optimal when there are n alterations, $PV_{III}(K',n') \geq PV_{III}(K,n')$. Thus, as claimed, $PV_{III}(K,n) < PV_{III}(K',n')$. *Q.E.D.*

B5. Proof of Proposition 5

It is enough to show that $p(K,n) \to 0$ as $n \to \infty$ (since there are finitely many possible values of K for every fixed N, the convergence is uniform). From Eq. (8), since $1 - \Phi\left(z(\alpha) - \kappa_i\sqrt{m}\right) \in [0,1]$

$$p_i(K,n) \leq \sum_{m=1}^{K/2} \Pr[2m \leq M \leq 2m+1|K,n] = \Pr[M \geq 2|K,n] \leq \frac{\mathrm{E}[M|K,n]}{2} = \frac{K}{2n} \to 0,$$

where the second inequality follows from Eq. (6) and Markov's inequality. *Q.E.D.*