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

PRELIMINARY MAJOR REVISION. DO NOT CIRCULATE

March 2, 2026

Abstract

Technological advances and genomic sequencing opened the road to personalized medicine: specialized therapies targeted to patients displaying specific molecular alterations. In lung cancer, targeted therapies are now available for 50% of patients—with some alterations affecting less than 1% of patients—greatly increasing life expectancy. In a model of entry and drug development, current institutions mandating approval of individual therapies imply lower or even null investments, higher prices, lower revenues, and fewer patients' enrollments in clinical studies for rarer alterations, predictions that are empirically supported. 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 spouse, 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.

"Bespoke, personalized therapies, which hold tremendous promise, challenge traditional models of drug and biologic development. The Food and Drug Administration (FDA) is committed to providing regulatory guidance and encouragement, outlining a path to market entry for products where a randomized trial is not feasible." Vinay Prasad (Chief Medical and Scientific Officer, FDA) and Martin A. Makary (Commissioner, FDA), New England Journal of Medicine, Nov 12, 2025.

"This month Britain's Medicines and Healthcare products Regulatory Agency (MHRA) approved a novel sort of clinical trial. Ten children, each suffering from an ultra-rare genetic neurodegenerative disease that threatens his or her life, will each receive a unique version of a known drug molecule. If the trial is successful, the MHRA will give the nod not to each custom drug one by one, but to the process of making them." The Economist, Jan 22, 2026.

1. Introduction

The publication of the initial sequencing of the human genome (International Human Genome Sequencing Consortium, 2001) opened the road to modern personalized medicine: that is, the delivery of the right drug to the right patient based on their genomic profile. Technological advances over the last twenty-five years have 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.

In this paper, we introduce a model of pharma companies' investment decisions, market entry decision, and bargaining between pharma companies and health insurance companies, to study the economic sustainability and the incentives for personalized drug discovery as scientific advances identify ever more specific molecular alterations that, on the 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) or the European Medicine Agency (EMA) 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. Our model incorporates the essential aspects of this process.

Given this structure, potential new firms must decide whether to enter the market to develop new drugs for alterations that affect small fractions of the lung-cancer population, anticipating that profits will depend on the pricing obtained through bargaining with health insurance companies, and that new competitors may enter the market. We show that (a) rarer alterations will elicit less or even no entry, and thus less or no new research; (b) the prices of drugs are higher for rarer alterations; (c) still, the total amount of revenues are lower for rarer alteration; (d) there are fewer patients enrolled in clinical trials for rarer alterations. We use data from lung-cancer research and empirically validate all these predictions.

Specifically, consistently with our theoretical predictions, data from lung-cancer treatments reveal fewer clinical trials for rarer lung-cancer alterations over the ten-year span since discovery of each alteration. Different horizons and specifications with year fixed effects provide similar results. Moreover, using current prices of best lung cancer treatments, we find that rarer alterations have far higher treatment costs. For instance, the best treatment of EGFR mutation (14.5% of lung cancer population) costs around $\$228,000$ per year, while the best treatment for the NTRK alteration (0.2% of lung cancer population) costs over $\$525,000$ per year. Alterations between these two extremes exhibit a log-linear relationship, with a regression R^2 of over 70%. But even with such higher treatment prices, we find that total revenues for pharma companies are higher for more common alterations, with a regression R^2 of over 98%. Finally, we extract the number of patients with lung cancer enrolled in clinical trials from published studies of approved treatments, and we find a positive non-linear relation between incidence and number of patients enrolled, as our model predicts.

[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.

Our model also predicts that economic incentives fully evaporate for the rarest alterations. We calibrate our equilibrium model to lung-cancer data to determine the incidence of alterations for which pharma companies will stop entering the market and thus investment in drug development will stall. We find that this cutoff is approximately 0.07% of the lung cancer population. We report in the next section that scientific research has already uncovered a large number of genetic alterations that affect *less* than 0.07% of the lung cancer population, thereby showing we are reaching the breaking point of the current approval regime. We emphasize that, because modern oncology increasingly treats lung cancer as a collection of numerous rare alterations, the overall development of new therapies for lung cancer will stall under the current regulatory framework.

If private pharma companies will not invest, it may in principle be feasible for governments to subsidize R&D. However, we show that the the societal costs of government-sponsored research would become unsustainable as the number of new rare alterations keeps increasing. This is due to the exorbitant R&D costs that even the government (the social planner) would need to bear in order to develop drugs for all rare alterations. We show that such costs would run in the trillions of dollars, and indeed, the cost per patient per year to treat such rare alterations would vastly exceed the current estimates of value per statistical life year. In other words, even from a societal perspective, it is not feasible to subsidize the research and the development of targeted drugs.

Our model, however, also suggests a pharmacological and regulatory path out of this impasse. Specifically, it indicates that a shift toward a framework in which regulatory agencies approve the *process* for drug discovery, rather than individual therapies, leads to a sustainable system that encourages R&D. Recent technological advances have made such a process realistic. In brief, a "personalized drug discovery process' (PDDP) would comprise five steps (see section 7.1. for details): (A) Identification of driver alterations, utilizing existing high-throughput genomic filters to pinpoint the specific mutation driving tumor growth; (B) Protein modeling: reconstructing the altered protein using advanced machine learning and AI, such as Google DeepMind AlphaFold3; (C) Compound matching: Leveraging supercomputers to screen billions of molecules to identify a compound that can "block" the altered protein and hence inhibit the growth of the tumor; (D) Safety and Absorption Profiling: Utilizing bioinformatics and molecular structure databases to ensure that the identified compound is safe, absorbable, and human-tolerable; (E) Drug production.

Current research in biochemistry and bioinformatics shows that steps (A) – (D) can now be performed in silico (i.e., through simulation) very rapidly, sometimes under one month (see e.g.Ren et al., 2023). Under this system, a patient is tested for alterations (step A); if a target is found, a personalized drug is developed through steps (B)–(D) and produced (step

E). If no new drug is identified, the patient defaults to existing standard-of-care therapies.

Ultimately, the PDDP model allows for the rapid identification of safe, personalized treatments in a fraction of the time required by traditional R&D and approval regimes. Most importantly, building on our calibration results, we show that the PDDP is economically sustainable. Because the regulatory approval cost is tied to the process rather than the individual alteration, it eliminates the need for prohibitively expensive RCTs for every alteration. We use the same model of investments, entry decision, and bargaining with insurance companies to estimate the drug prices that would be charged by pharma companies under the PDDP model. We estimate drug prices in the range of \$250,000-\$275,000, depending on the ability of the PDDP to find new alterations and new therapies, and the survival rates from such therapies. Moreover, and importantly, the present value of future profits increases substantially as the PDDP becomes more effective in detecting new molecular alterations and increasing the survival rate of patients. The critical point is that PDDP applies to all lung-cancer patients rather than subgroups; therefore, profits scale up with improvements in detection and survival probabilities. This in turn implies that pharma companies retain the incentive to develop more effective PDDPs—which benefits both the companies themselves and their patients. The latter would receive highly personalized treatments that would increase their life expectancy at costs that are not as exorbitant as the current ones.

Although approving a process for drug discovery as opposed to individual drugs would be a significant departure from current regulatory practices, the quotes at the beginning of this article demonstrate that both the FDA in U.S. and the MHRA in U.K. are warming up to the idea of embracing this different paradigm, once it is supported by a full RCT to validate the process, as we discuss in Section 7.2.

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.

While these contributions focus on different aspects and trade-offs related to experimentation and discovery of new, safe treatments, we take the discovery of new targetable alterations as the starting point, take the regulatory process as given, and study the resulting incentives for pharma companies to enter the market over time. Furthermore, we compare the current regulatory regime with one that is tailored to personalized medicine.

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. Furthermore, we analyze the entry of new firms in the market for (targeted) treatments.

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

[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.

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..

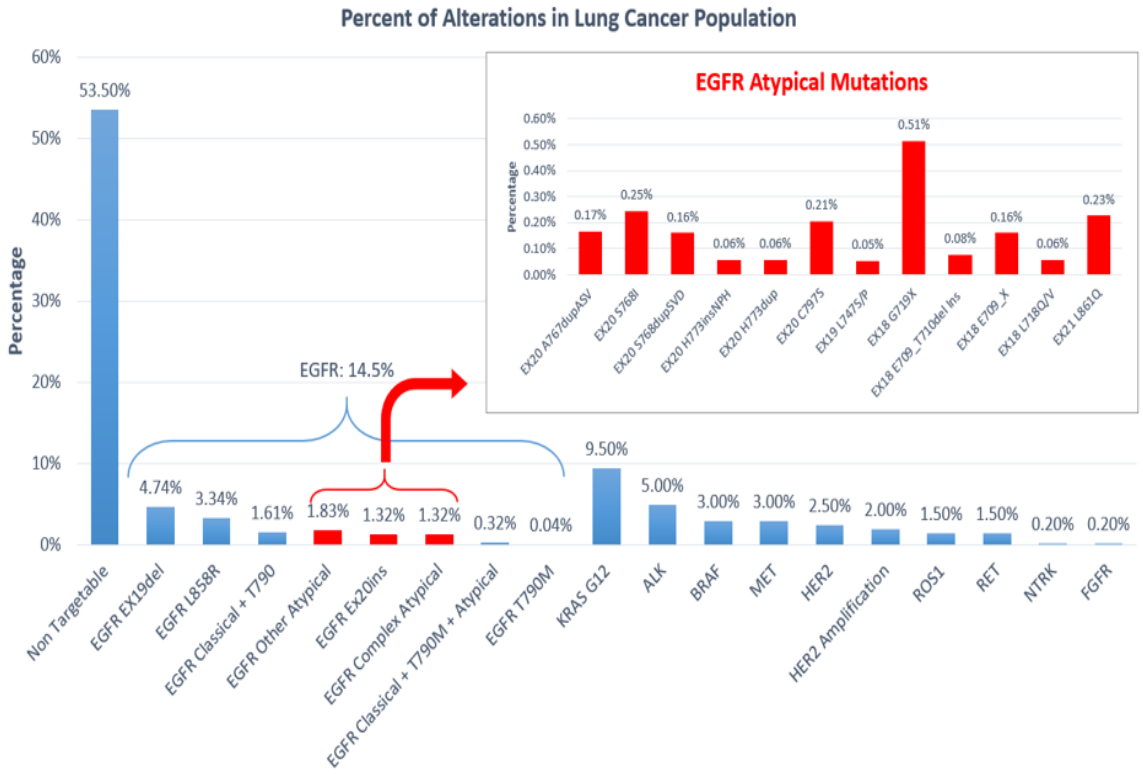
Following the discovery of EGFR mutations in 2004, several other targets were identified. ALK and ROS1 fusions were identified in 2007, followed by BRAF, RET, HER2, MET, NTRK, KRAS G12, EGFR Exon 20, and NRG1. Figure 1 indicates the frequency of alterations in lung cancer patients.[3] About 53.5% of 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 0.2% 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 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 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

[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)

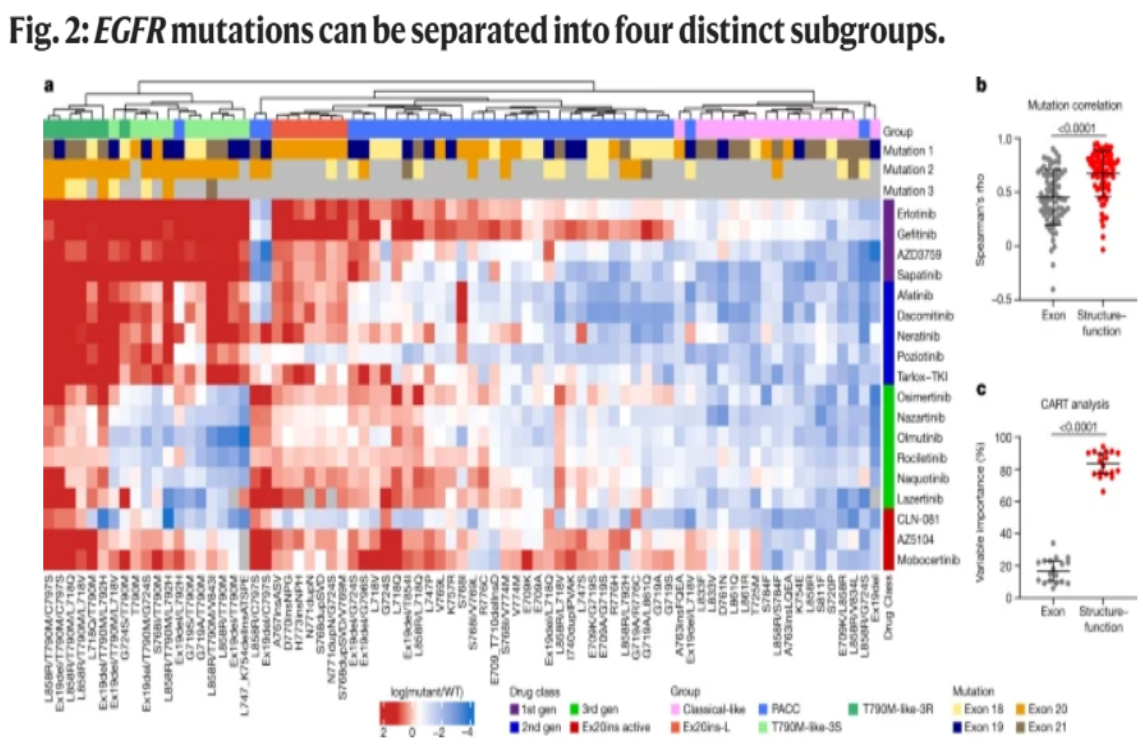
drugs has not been confined to the treatment of lung cancer: it has extended to many other tumors, such as melanoma, thyroid cancer, gastric cancer, etc. In addition, over time, these tumors have become resistant to the first generation of drugs. Consequently, new generations of drugs have been developed 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.

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,

[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.

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

Figure 2: **EGFR alterations**



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

Figure 3: **Drug Discovery Progression**

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.

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*	$1,942	10.8 vs 5.4	
Erlotinib	vs. chemo	Roche	2004*	$2,078	9.7 vs. 5.2	$124
Afatinib	vs. chemo	Boehringer	2013	$2,939	12.4 vs. 6.8	
Dacomitinib	vs. Gefitinib	Pfizer	2018	$4,115	14.7 vs. 9.2	
Osimertinib	vs. Gefitinib	AstraZeneca	2018	$4,394	18.9 vs. 10.2	$594
ALK (3% - 7%)						
Crizotinib	vs. chemo	Pfizer	2011	$5,924	7.7 vs. 3.0	
Ceritinib	vs. chemo	Novartis	2017	$3,681	16.6 vs. 8.1	$1,095
Alectinib	vs. Crizotinib	Roche	2015	$4,851	34.8 vs. 10.9	$409
Brigatinib	vs. Crizotinib	Ariad	2020	$5,181	24.0 vs. 11.1	$491
Lorlatinib	vs. Crizotinib	Pfizer	2021	$5,322	NR vs. 9.1	
Ensartinib	vs. Crizotinib	Xcovery	2024	$4,781	25.8 vs. 12.7	
ROS1 (1% - 2%)						
Crizotinib	vs. chemo	Pfizer	2016	$5,924	18.4 vs. 8.6	
Entrectinib	SA	Roche	2019	$3,681	19.0	$451
Repotrectinib	SA	Brystol Myers Squibb	2023	$7,667	35.7	
RET (1% - 2%)						
Selpercatinib	vs. chemo	Eli Lilly	2022	$5,914	24.8 vs. 11.2	$511
Pralsetinib	SA	Roche/Blueprint	2023	$5,556	NR	$260
BRAF (1% - 5%)						
Dabrafenib tramentinib	SA	Novartis	2017	$6,217	14.4	$1,239[1]
Encorafenib Binimetinib	SA	Pfizer	2023	$7,402	30.2	
MET ex 14 (2% - 4%)						
Capmatinib	SA	Novartis	2020	$6,232	12.4	$561
Tepotinib	SA	Merck	2021	$6,423	11.0	
NTRK (<1%)						
Larotrectinib	SA	Bayer	2018	$10,101	24.6	$809
Entrectinib	SA	Roche	2019	$5,244	11.2	$451
EGFR Exon 20 (1.3%)						
Amivantamab - chemo	vs. chemo	J&J	2021*	$7,709	11.4 vs. 6.7	
Panel B: Second Line						
EGFR (12% - 17%)						
Osimertinib	vs. chemo	AstraZeneca	2018	$4,394	10.1 vs. 4.4	$594
Amivantamab/chemo	vs. chemo	J&J	2021	$7,709	8.2 vs. 4.2	
KRAS G12C (7% - 12%)						
Sotorasib	SA	Amgen	2021	$5,419	6.8	
Adagrasib	SA	Mirati	2022	$5,629	6.5	
HER2 (1% - 4%)						
Trastuzumab Deruxtecan	SA	Roche	2022	$4,943	8.2	

Source: National Cancer Institute, FDA, Drugs.com, Taha et al. (2021). Prices are from drug.com, accessed in 5/28/2025. 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.

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, 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

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

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

3.1. Overview

Our stylized market for treatments features four different types of participants: patients, pharma companies, insurance companies, and the regulatory authority. Patients do not have an active role in our model. As far as the regulator is concerned, we do not take a mechanism-design perspective and instead study the implications of the current regulatory regime, as well as (in Section 7.2.) of an alternative regime that seeks to approve drug development processes rather than individual drugs.

The main agents in our model are firms and insurers. Section 3.2. outlines the timeline of the R&D and approval process needed before a pharma company is able to start marketing a drug. Sections 3.4.–3.8. analyze the decisions faced by a pharma company in the course of the approval process, highlighting the (lack of) incentive to pursue approval for treatments targeting rare alterations or diseases. Throughout these sections, firms take drug prices as given. Sections 3.9. and 3.10. close the model, assuming that prices are determined by bargaining between pharma companies and insurers, and solving for the entry decision of new pharma companies. Finally Sections 3.11. and 3.12. show that, under the assumed regulatory regime and market structure, the pharma industry as a whole will not have sufficient incentives to develop treatments for rare alterations; furthermore, if a social planner wanted to subsidize R&D for such rare alterations, the cost would be prohibitive—in fact, unbounded as the number of alterations grows.

We explicitly model the multi-sided nature of markets for drugs for two reasons. First, in most developed economies, targeted treatments are partially paid for by intermediaries (whether insurance companies or the national health service). Second, while stylized, our model is still rich enough to match many empirical facts on drug development and pricing for rare targets: we demonstrate this in Section 4.. The richness of the model allows us to calibrate the model and provide forecasts about pricing even in a new paradigm in which regulatory agencies approve the process, rather than individual therapies, as developed in Section 7.1.

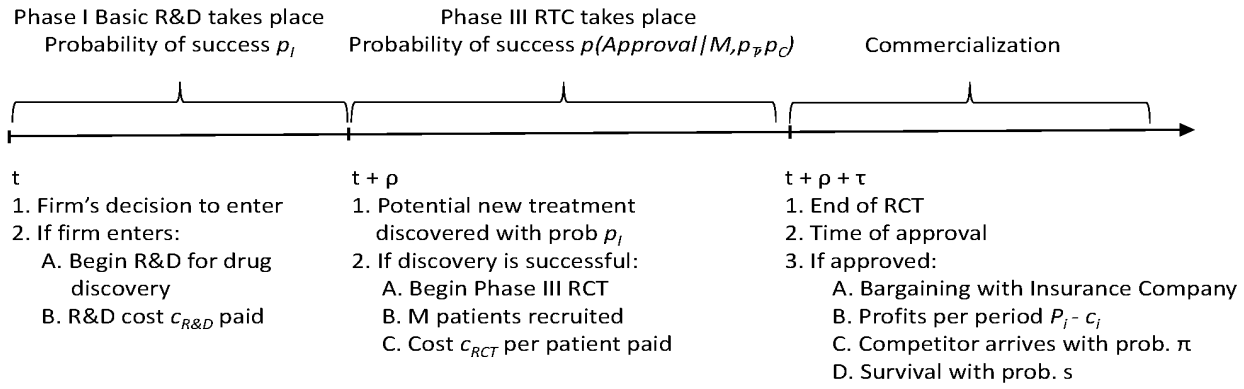
3.2. Drug development and approval

Let N be the number of new patients diagnosed with a given illness per year. At time $t = 0$, a potential new alteration is discovered. In every period $t \geq 0$, a new pharma company arrives with probability $\mu \in [0, 1]$ and must decide whether to invest to start developing a new drug specific for that alteration. Developing a new 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 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 5.) 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 5.. 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.

[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



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.

3.3. Approval probabilities for given subject pool

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 $H_0 : p_T \leq p_C$.

[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.

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.

3.4. Pharma Profits and Discounting

Let P denote the expected price of the new treatment and c the marginal cost of producing it, once it has been approved. The profit to the pharma company per year and per individual is $(P - c)$ 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 $N \times (1 + s + s^2 + .. + s^j) = N \times \frac{1-s^j}{1-s}$ 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 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 $(1-\pi)^{\tau+j}$. We endogenize π in Section 3.10.. 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

$$PV_{III} = A_{III} \cdot (P - c) \cdot N,$$

where A_{III} is the after-tax discount factor

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

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

[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.

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.

3.5. Personalized Medicine: finding a large enough subject pool

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. Let each subgroup i have size $q_i \in (0, 1)$, with $\sum_{i=1}^{n} q_i = 1$ (see Figure 1). 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

$$\Pr(M|K, q_i) = \begin{pmatrix} K \\ M \end{pmatrix} (q_i)^M (1 - q_i)^{K-M} \tag{5}$$

Hence, alterations affect a smaller size of the population with a given disease, 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,

$$E[M|K, q_i] = Kq_i \tag{6}$$

As q_i decreases, for a fixed total number K of subjects recruited, the expected number of patients in subgroup i 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 for group i will only be beneficial to Nq_i patients, rather than the full group of N patients. To sum up, as the fraction q_i decreases, it becomes harder to recruit a sufficient number of patients for each subgroup i, *and* the addressable market for each targeted therapy also shrinks.

3.6. Phase III for personalized medicine

From sections 3.4., the present value formula for each treatment i with incidence q_i is :

$$PV_{III,i}(K, q_i) = \{p_i(K, q_i)\, A_{III}\ (P_i - c_i)\ N\ - c_{RCT}\, K\}\ q_i \tag{7}$$

where A_{III} is as in Eq. (4) and $p_i(K, q_i)$ it the probability of approval, given by

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

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.

3.7. R&D Disincentives for Rarer Targets

We now assume that q_i is sufficiently small so that the approximate variance of M, namely $Kq_i(1-q_i)$, is approximately Kq_i:

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

$$M \sim \phi(M, Kq_i, Kq_i). \tag{9}$$

We then obtain:

Proposition 1 *Let the Normal Approximation Assumption hold. Then, the probability of approval $p_i(K, q_i)$ is increasing in q_i and K:*

$$\frac{\partial p_i(K, q_i)}{\partial q_i} > 0; \qquad \frac{\partial p_i(K, q_i)}{\partial K} > 0 \tag{10}$$

That is, as the medical profession identifies targets that affect fewer patients, 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 2 *Let the Normal Approximation Assumption hold. Then, everything else equal, the present value $PV_{III,i}(K, q_i)$ decreases as q_i decreases.*

Clearly, a Pharma company can take two actions as a consequence of proposition 2: Increase the number of patients K and/or increase the price P_i. We study the latter strategy further below. As for the former strategy, from Equation (7), increasing K only works if $\partial p(L, q_i)/\partial K$ is sufficiently large. Moreover, K is bounded above by N itself. The following proposition, which does not rely on the normality assumption, makes the point that such strategy does not work for small q_i:

Proposition 3 *For every value of A_{III}, P_i, c_i and N, there is $\bar{q}_i$ such that $PV_{III}(K, q_i) < 0$ for all $q_i \leq \bar{q}_i$.*

Proposition 3 captures a straightforward, yet important insight: q_i is small, then it will be virtually impossible for pharma companies to conduct an RCT, simply because, for any alteration, there will not be sufficient patients in the treatment group. Under the current approval process in which regulatory agencies approve individual therapies, a truly personalized medicine is economically infeasible.

3.8. Phase I for personalized medicine

The Pharma company will take into account the difficulties to secure a proper sample size earlier on in the development of the new drugs. Indeed, moving backward to time 0, 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 $p_{I,i}$ be the probability of advancing to Phase III (i.e. to the randomized clinical trial), and denote by

$$A_I = \frac{(1-\pi)^\rho}{(1+R)^\rho} \tag{11}$$

the discount factor corresponding to Phase I, which takes into account the time ρ between Phase I and Phase III, and the fact that a competitor can arrive before τ to steal the market. Then, the 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

$$PV_I(q_i) = p_{I,i}\, A_I\, \{p(K^*, q_i)\, A_{III}\, (P_i - c_i)\, N - c_{RCT}\, K^*\}\, q_i - c_{R\&D,i} \tag{12}$$

where K^* is the optimized number of patients performed in previous step (Phase III). Clearly, as q_i decline, so does $PV_{I,i}$, which reaches zero for a higher value of q_i than $PV_{III,i}$ because of the additional discount $p_{0,i}\, A_0$ and the additional R&D cost $c_{R\&D,i}$. A result similar to Proposition 3 clearly holds in this case as well.

3.9. Insurance Market and Bargaining

We now endogenize the drug price P_i through a simple bargaining model between pharma companies and health insurance companies. We only consider insurance for a single disease of interest (e.g. lung cancer) to maintain a sharp focus. Let the total insured population be N_{pop}, and let q be the i.i.d. probability that any insured individual develops the disease. Each person pays a health insurance premium C; if the person is affected by the disease, the insurance company fully covers the treatment (so there are no copays or deductibles). The total profit for the insurance company is thus

$$\Pi = C\, N_{pop} - N_{pop}\, q \sum_{i=1}^{n} P_i\, q_i \tag{13}$$

For each alteration i the insurance company sets a maximum price $\overline{P}_i$, such that, at these prices, the insurance company would break even in expectation:

$$\frac{C}{q} = \sum_{i=1}^{n} \overline{P}_i\, q_i \tag{14}$$

Thus, the profit for the insurance company is

$$\Pi = N \left(\sum_{i=1}^{n} (\overline{P}_i - P_i)\, q_i \right) \tag{15}$$

where $N = q\, N_{pop}$ is the number of patients per year with the diseases, and P_i is the *actual* price the insurance company and the firm agree upon, as we describe momentarily. Each maximum price $\overline{P}_i$ for therapy i may depend on i, as well as on the fractions $q_1, \ldots, q_n$. While we do not micro-found the determination of such maximum prices, we estimate them below from observed prices.

The actual price P_i of each drug is set by bargaining between each insurance company and the pharma company. The pharma company will not enter the market and develop the new drug if the present value PV_I in equation (12) is negative. Solving for the break-even price, the pharma company will not enter the market if

$$P_i < P_i^{BE} = c_i + \frac{p_I\, A_I\, c_{RCT}\, K\, q_i + c_{R\&D}}{p_I\, A_I\, p(K, q_i)\, A_{III}(q_i)\, N\, q_i}. \tag{16}$$

If the pharma company had zero bargaining power, then $P_i = P_i^{BE}$ for all i.

If the pharma company had full bargaining power, instead, then $P_i = \overline{P}_i$: that is, the drug price would exactly match the maximum price the insurance company is willing to pay, and the insurance company ends up with zero profits. We assume a simple α-split in the bargaining power, so that the drug price P_i is set according to

$$P_i = \alpha\, \overline{P}_i + (1 - \alpha)\, P_i^{BE} \tag{17}$$

under the condition $P_i \leq \overline{P}_i$.

Proposition 4 *Everything else equal, the breakeven price P_i^{BE} increases as q_i declines. Assuming the max price $\overline{P}_i$ does not decline faster than q_i as q_i declines, the price P_i for alteration i increases as q_i declines. Finally, in the limit, the expected payment for alteration i, $P_i q_i$ declines as q_i declines.*

There are two reasons why the break-even price P_i^{BE} increases as q_i declines, when $c_{R\&D} > 0$: In the denominator, both the incidence q_i and the probability of approval $p(K, q_i)$ decline as the incidence q_i declines (cf. Proposition 1). Moreover, while the former effect disappear in the (unrealistic) case of $c_{R\&D} = 0$, the latter remains nonetheless, showing that the break-even price increases even when R&D costs are negligible.

Proposition 4 implies that everything else equal, we should see higher prices for alterations that are less common. However, even so, the expected payment per alteration declines as q_i declines in the limit, because the maximum prices $\overline{P}_i$ are finite and the constraint $P_i < \overline{P}_i$ must hold. In the calibration section, we estimate the maximum prices $\overline{P}_i$ from the data under two simple α-splits between the pharma company and the insurance company.

3.10. Endogenous Entry and Intensity of R&D

Finally, we close the model by analyzing the equilibrium entry decision of pharma companies. This also allows to analyze the incentives for drug development when other firms can enter the market. Specifically, we now endogenize the probability π that a competitor enters the market with a superior treatment and steals the market from the incumbent. We make the simplifying assumption that the decision time t is sufficiently large so that the model is (nearly) stationary.[11]

We assume that firms have heterogeneous R&D costs $c_{R\&D}$, independently drawn from the uniform distribution on some interval $[c_L, c_H]$. Each firm knows its own R&D cost $c_{R\&D}$, but is uncertain about its standardized improvement coefficient κ; specifically, every firm knows that κ is independently drawn according to the same density $f_\kappa(\cdot)$.

The entry equilibrium works as follows: In every period there is probability $\mu(q_i)$ of a new firm arriving to the market; the firm must decide whether to enter the market for alteration i or not; this probability $\mu(q_i)$ determines the probability $\pi(q_i)$ of successful entry in every period; hence, the discount factors $A_I(q_i)$ and $A_{III}(q_i)$ are computed using this value $\pi(q_i)$

[11]The non-stationarity of the model for low t is evident at time $t = 0$ and $t = 1$. At time $t = 0$, no competitor will enter the market for τ periods, while at time $t = 1$ no competitor will enter the market for $\tau - 1$ periods. A firm's decision problem at these two times are thus (slightly) different. When t is large, a competitor can endogenously always enter at any time and the decision problem at any t is always the same.

as the per-period probability of an entrant taking away the market; in turn, the discount factor determines the present value of entry, which then identifies the probability $\mu(q_i)$ itself that the firm has $R\&D$ cost $c_{R\&D}$ low enough to make entry profitable.

To analyze this situation, from equation (12), rewrite the present-value formula as

$$PV_I(q_i) = p_{I,i}\, A_I(q_i)\, \{p(q_i)\, A_{III}(q_i)\, (P_i - c_i)\, N - c_{RCT}\, K^*\}\, q_i - c_{R\&D,i} \tag{18}$$

where P_i is the price determined in the previous section, and the probability $p(q_i) = p_i(K^*, q_i)$ is as in (8) except with $E_\kappa\left[\Phi\left(z(\alpha) - \kappa\sqrt{m}\right)\right]$ in its last term. We now denote the discount factors $A_I(q_i)$ and $A_{III}(q_i)$ to make the dependence on the incidence q_i of alteration i explicit. The reason is that the probability of successful entry, denoted π in the discount factor in equation (4), is now endogenously given by

$$\pi(q_i) = p_{I,i}\, p(q_i)\, \mu(q_i) \tag{19}$$

where $p_{I,i}$ and $p(q_i)$ are the probability of success of phase I and phase III (the latter is endogenous), and $\mu(q_i)$ is the probability that the arriving firm decides to enter the market to develop the new drug, i.e., such that its $PV_I(q_i) \geq 0$. The latter type of firms are those with sufficiently low costs of $R\&D$. Recalling the assumption $c_{R\&D} \sim U[c_L, c_H]$, conclude that $PV_I(q_i) > 0$ if (a) $PV_{III}(q_i) = p(q_i)\, A_{III}(q_i)\, (P_i - c_i)\, N - c_{RCT} K > 0$; and (b) $c_{R\&D} < \bar{c}_{R\&D}(q_i)$ where the cutoff $\bar{c}_{R\&D}(q_i)$ is given by

$$\bar{c}_{R\&D}(q_i) = p_I\, A_I(q_i)\, \{p(q_i) A_{III}(q_i)(P_i - c_i)\, N - c_{RCT}\, K^*\}\, q_i. \tag{20}$$

In this equation, we also recall that the price P_i depends itself on the R&D costs $c_{R\&D}$, as firms with higher R&D costs have a higher break-even prices and thus are able to get higher prices through bargaining (see equations (17) and (16)). We show in the calibration section how to estimate this cut-off from observed drug prices.

To sum up, the probability $\pi(q_i)$ of competing firms entering the market depends on the probabaility $\mu(q_i)$ that an arriving firm has $PV_I(q_i) > 0$, where the latter depends on the probability $\pi(q_i)$ through its impact on the discount factor $A_{III}(q_i)$. The following proposition establishes the existence of a fixed point $\mu^*(q_i) \to \pi^*(q_i) \to \mu^*(q_i)$:

Proposition 5 *If $c_L > 0$ is sufficiently low, and if $PV_{III}(q_i) > 0$, there is a fixed point $\mu^*(q_i) = \frac{\bar{c}^*_{R\&D}(q_i) - c_L}{c_H - c_L}$ such that $\bar{c}^*_{R\&D}(q_i)$ is given in (23) where the discount factors $A_I(q)$ and $A_{III}(q)$ are calculated with*

$$\pi^*(q_i) = p_{I,i}\, p(q_i)\, \mu^*(q_i) \tag{21}$$

The following result is immediate:

Proposition 6 *If $c_L > 0$, then there is q_i sufficiently low for which $\bar{c}_{R\&D}(q_i) < c_L$ and thus $\mu^*(q_i) = 0$ is a fixed point and no entry occurs in steady state.*

The last two propositions show that there are two values of q_i such that the probability of a potential entrant is larger for higher q_i. While a more general statement is harder to establish due to the non-linearity of $PV_{I,i}(q_i)$, a continuity argument suggests a positive relation between q_i and $\mu^*(q_i)$ for some range of q_i. That is, alterations with higher incidence in the population attract more investments in drug discovery. We will find such relation in our calibrated simulations below.

3.11. R&D Disincentives at the Industry Level

As a final point, it is useful to aggregate the previous results at the Pharma industry level. In particular, we analyze research incentives when the number of alterations keeps increasing and therefore each affects an increasingly smaller fraction of the patient population. In this section, we then assume that there are n alterations and all alterations are symmetric, so that for all $i = 1, ..., n$ we have $q_i = 1/n$. Correspondingly, we assume that the price P_i is the same for each alteration i, and denote this common price by $P(n)$ to emphasize the dependence on the number of alterations. Similarly, we assume that all firms have the same marginal cost c_p and standardized improvement κ. The aggregate present value of developing treatments for all alterations depends on the specific R&D costs of the pharma companies that are targeting each alteration $i = 1, .., n$. Recalling that such $c_{R\&D} \sim U[c_L, \bar{c}_{R\&D}]$, the aggregate expected present value at the industry level using the expected R&D cost:

$$E[PV_I(n)] \;=\; p_I A_I(n) \left\{ p(K^*(n), n)\, A_{III}(n)\, \left[P(n) - c_p\right] N \, - c_{RCT}\, K^*(n) \right\} - 0.5(c_L + \bar{c}_{R\&D})\, n \tag{22}$$

where the acceptance probability $p(K^*(n), n)$ is as in Equation (8) with $q_i = 1/n$, and $K^*(n)$ reflects the optimized number of patients for each n. The discount factors also depend on n as the equilibrium incentives to enter the market depends on $q_i = 1/n$, as discussed in previous section.

As the number n of discovered targeted therapies increases, the incidence $q = 1/n$ for each pharma company decreases. Our earlier equilibrium results (cf. Proposition 4) suggest that both the price $P(n)$ and discounts $A_I(n)$ and $A_{III}(n)$ thus increase as n increases. On the other hand $p(K^*(n), n)$ decreases for n large, and the last term $0.5(c_L + \bar{c}_{R\&D})\, n$ also causes $E[PV_I(n)]$ to decrease as n increases. The outcome of these conflicting forces cannot be pinned down without additional assumptions. The calibration section shows that we in fact obtain a hump shape for $E[PV_I(n)]$: as n increases, the effect on prices and discount

initially dominate, and for n large, the declining $p(K^*(n), n)$ and the last term dominate, thereby pushing $E[PV_I(n)]$ to zero. Indeed, we can show a general limiting result for n large (this does not depend on the normal approximation assumption):

Proposition 7 *For every composition of parameter values, there is $\bar{n}$ such that $E[PV_I(K, n)] < 0$ for all $n \geq \bar{n}$.*

Proposition 7 shows that, as the number of alterations keeps increasing, the present value of future profits at the industry level will become negative. This implies that even aggregating across all alterations, it will become unprofitable to conduct research on targeted therapies: under the current approval process in which regulatory agencies approve individual treatments, truly personalized medicine is economically infeasible.

3.12. The Social Costs of Personalized Medicine

The analysis above shows that even even science generates significant improvements in healthcare for a given disease via personalized treatments, it is effectively impossible for the whole industry to be profitable. In turn, this will stall scientific research. In principle, the government could step in and sustain science development. However, we now show that the social costs for government intervention are unbounded as the number of new discoveries n increase.

Consider a social planner who wishes to develop targeted treatments for all alterations, but must nevertheless conduct clinical trials as in the current approval process; this seems both realistic and necessary to ensure that patients and doctors trust the outcomes of scientific research. To keep the analysis consistent with the previous section, we calculate the breakeven price P^{BE} that ensures that $E[PV_I] = 0$. This yields the minimum annual cost of treating each patient, which can be directly compared with standard measures of value per statistical life year. We set the corporate tax t_c and any tax credit incentives to zero: those represent transfers to and from the social planner, which can be omitted if the social planner herself sets prices. Under these assumptions, the break-even price is the same as in equation (16) with $q_i = 1/n$.

The following proposition summarizes results from the preceding sections, under the parametric assumptions just described.

Proposition 8 *For every value of $c_p, c_{R\&D}$ and N, and for every P^*, there is $\overline{n}$ such that $P^{BE} > P^*$ for $n > \overline{n}$.*

Proposition 8 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. (16) 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 $P^{BE} \times N/(1-s_T)$ where $N/(1-s_T)$ is the average number of patients treated per year in steady state.

Unbounded societal costs occur even if the *R&D* cost of Phase I, $c_{R\&D}$, is zero: the driving factor in this case 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.

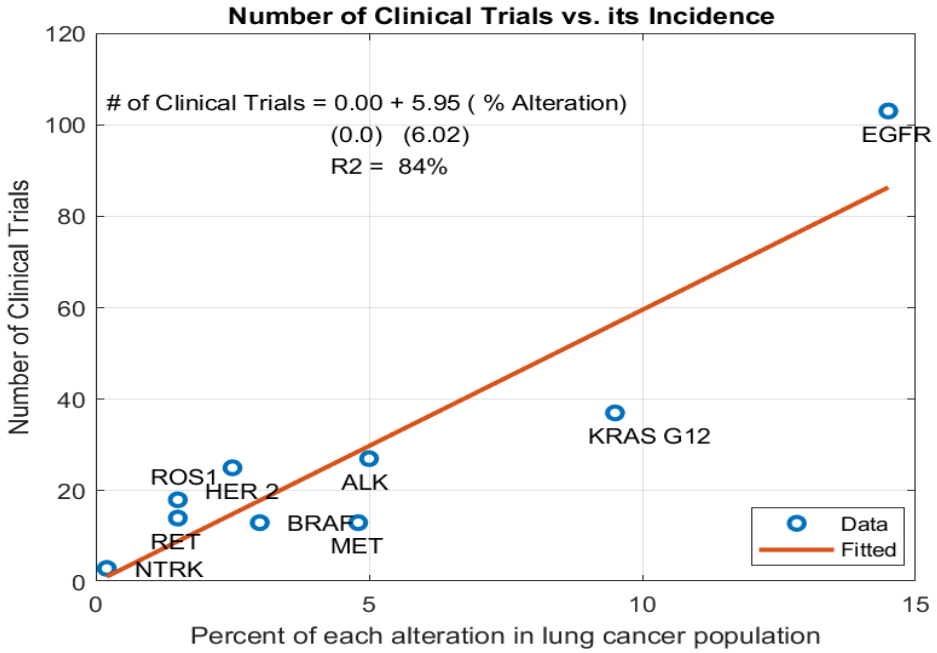
4. Model Predictions and Evidence

To summarize, our model predicts that even within the same disease, such as lung cancer, (1) rarer alterations are subject to less research from pharma companies (Proposition 6); (2) rarer alterations have higher drug prices (Proposition 4); (3) rarer alterations have a lower average revenues (Proposition 4.); (4) rarer alterations have lower enrollment in clinical trials.

Prediction (1) implies that, as the discovery of alterations opens the door to personalized targeted therapies, economic incentives will prove to be an obstacle to the development of treatments for such alterations—as indeed is the case for rare diseases. We now provide evidence on this prediction.

Figure 5 shows the relation between the cumulative number of interventional 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 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

Figure 5: **Cumulative Number of Clinical Trials vs. its Percentage in Population**



Notes: The figure plots the cumulative number of clinical trials over a 10-year period since discovery of each alteration versus its incidence in the lung cancer population. 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. T-statistics are in parenthesis.

0.5% of the population, only 3 clinical trials have been carried out.[12] 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.

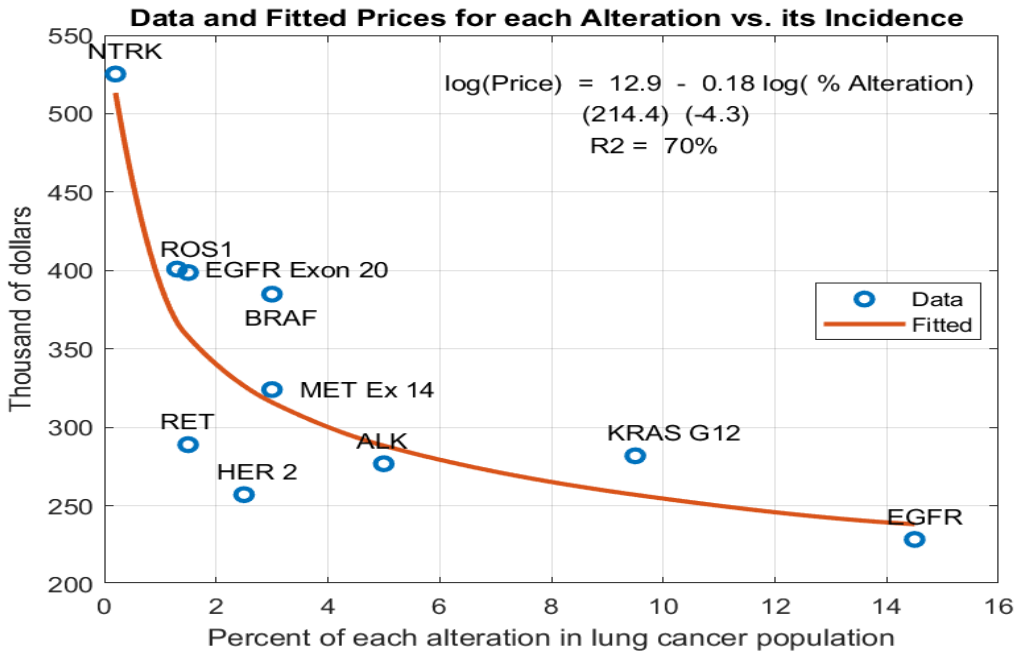
Turning to drug prices, we first collect the price for the best available treatment for each alteration, according to the data in Table 1. Figure 6 shows the data and the fitted line from the simple regression

$$log(Price_i) = \alpha + \beta\, log(Incidence_i) + \epsilon_i$$

We find strongly significant coefficients with an $R^2 = 70\%$, showing that in the data, indeed, lower incidence correlates with higher prices, as the model predicts. A caveat is that the source of the price data is drugs.com, which reports retail drug prices and not those nego-

[12]We also included time-fixed effects to ensure the result is not due to a timing effect, e.g. more biotech research in earlier periods than later periods, and the results are the same. Similarly, similar results obtain if we consider a shorter horizon than ten years.

Figure 6: **Prices per Year of Best Treatment for each Alteration versus its Incidence**



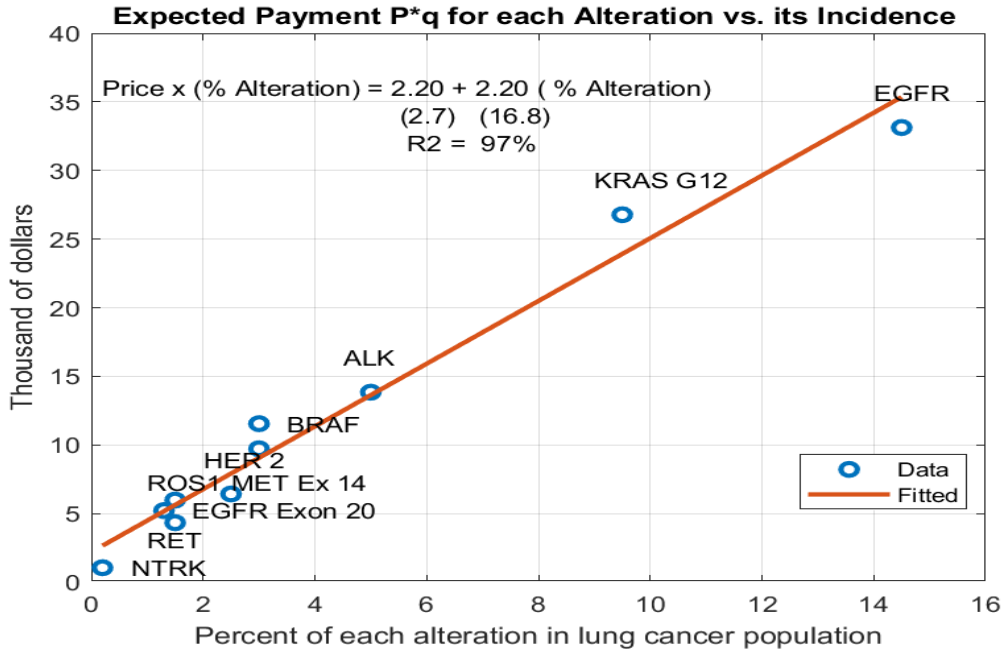
Notes: Price data are from Drugs.com and reported in Table 1 and they are for Osimertinib (EGFR), Lorlatinib (ALK), Repretrectinib (ROS1), Pralsetinib (RET), Enocorafenib plus Binimetinib (BRAF), Capmatinib (MET Ex 14), Larotrectinib (NTRK), Amivantanab (EGFR Exon 20), Sotorasib (KRAS G12), and Trastusumab Deruxtecan (HER 2). The percentage of lung cancer patients affected by each alterations on the horizontal axis is from Figure 1.

tiated with insurance companies. However, assuming that negotiation has a similar impact on prices across alterations, the empirical results would still hold.

Figure 7 shows that the expected revenue per patient per year is linearly related to its incidence, with an $R^2 = 97\%$. That is, while prices are indeed higher for less common alterations, the total average revenue per alteration is lower for the same, given their lower incidence.

Finally, Figure 8 shows the average enrollment of lung cancer patients that are reported in the published studies underlying the statistics in Table 1 for each alteration versus its incidence in the lung cancer population. Although data are more scattered in this case, we can find a non-linear positive relation between the two quantities, Although the choice of sample size certainly depends on many factors, including the expected standardized improvement of each therapy over standard of care, the positive relation also suggests more difficulty in securing large samples for more rare alterations.

Figure 7: **Expected Revenues Pq of Best Treatment vs. its Incidence**



Notes: Price data are from Drugs.com and reported in Table 1 and they are for Osimertinib (EGFR), Lorlatinib (ALK), Repretrectinib (ROS1), Pralsetinib (RET), Encorafenib plus Binimetinib (BRAF), Capmatinib (MET Ex 14), Larotrectinib (NTRK), Amivantanab (EGFR Exon 20), Sotorasib (KRAS G12), and Trastusumab Deruxtecan (HER 2). The percentage of lung cancer patients affected by each alterations on the horizontal axis is from Figure 1.

5. Calibration to Lung Cancer Research

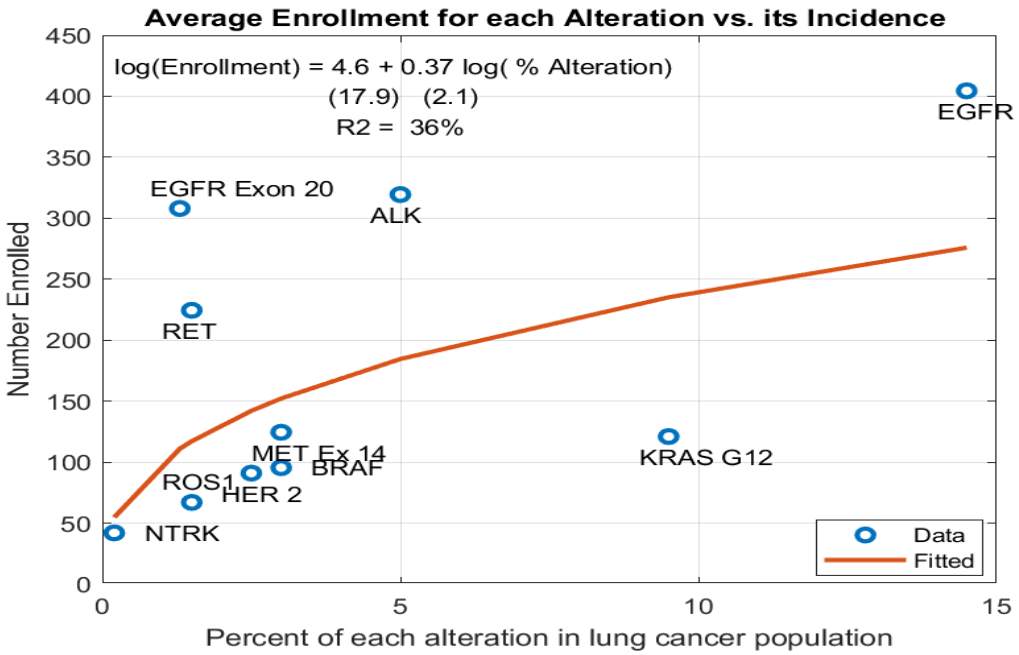
The previous sections shows that as researchers discover new alterations that affect an increasingly smaller fraction of the population, the approval probability drops and the financial incentives to undertake R&D investments evaporate. The question is for what fractions q_i such profits vanish and what are the social costs of personalized medicine under the current approval framework. In this section we provide a calibration using parameters from lung cancer research to answer these questions.

5.1. Parameters

Table 2 collects the parameters of the calibration. First, there are about 2.5 million new lung cancer cases per year in the world, including about 220,000 in the U.S.[13] However, 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. Thus, we set the number of new

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

Figure 8: **Average Patients' Enrollment of Approved Treatments for each Alteration vs. its Incidence**



Notes: Lung cancer enrollment data are from the published articles for each of the studies in Table 1. The percentage of lung cancer patients affected by each alterations on the horizontal axis is from Figure 1.

targetable cases per year at $N = 2.5$ million/2=1.25 million. The population that could be subject to clinical trials, however, is lower, due to eligibility criteria. We assume that only 10% of potential patients are eligible to be included in a clinical study (see e.g. Unger et al., 2024).

The length of a randomized clinical trial is set at $\tau = 4$ years, while the pre-clinical, Phase I and Phase II are set to $\rho = 8$ years. As for the success rate of the new treatment, we choose parameters based on the Osimertinib against Gefitinib trial in EGFR. In this case, we have $\kappa = 0.344$. We assume that all entrant firms have an expected standardized improvement $E[\kappa] = 0.35$, but also a 30% probability that $\kappa < 0$—an optimistic number, since about 36% of clinical trials with biomarkers in oncology fail in Phase III (see Wong et al., 2019, Table 3). Under the uniform distribution assumption, $\kappa \sim Uniform[a, b]$, with $a = -.525$ and $b = 1.225$, which yields the desired properties. We use Monte Carlo simulation together with small-sample statistics to ensure the proper calculation of rejections probability when sample sizes are small, since the normal distribution assumption discussed in Section 3. is only appropriate for large sample sizes. The appendix discusses the nature of the Monte Carlo simulations.

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. Finally, we set the R&D cost pre-Phase III to between \$50 million and \$1 billion, consistent with the costs calculated from Henderson et al. (2023, Supplemental Table S1) and reported in Table 1.

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.[14]

Finally, we use observed drug prices $\widehat{P}_i$ to estimate the cut-off $\bar{c}_{R\&D}(q_i)$. In particular, the bargained price (17) depends on the pharma company's R&D cost $c_{R\&D}$ through its break-even price (see formula (16)). According to the model, the observed price must depend on that firm's realized value of $c_{R\&D}$, which is not observable. Because we cannot observe the insurance company's maximum price $\overline{P}_i$ either, we cannot identify $\bar{c}_{R\&D}$. However, we can exploit the distribution over $c_{R\&D}$ conditional on $PV_I > 0$ – a uniform distribution between c_L and $\bar{c}_{R\&D}$ – and assume the current observed price is determined through bargaining between the insurance company and the average pharma company with $c_{R\&D}$ given by its expected value: $E[c_{R\&D}] = 0.5(c_L + \bar{c}_{R\&D})$. We thus obtain:

Proposition 9 *Let the observed price $\widehat{P}_i$ be determined by bargaining between the insurance company and the pharma company with expected R&D costs $E[c_{R\&D}] = 0.5\,(c_L, \bar{c}_{R\&D}(q_i))$. The cut-off $\bar{c}_{R\&D}(q_i)$ is then given by*

$$\bar{c}_{R\&D}(q_i) = \frac{2}{1+\alpha}\, p_I\, A_I \left\{ p(q_i) A_{III}(q_i) \widehat{P}_i (1 - pc)\, N - c_{RCT}\, K^* \right\} q_i - \frac{1-\alpha}{1+\alpha}\, c_L \tag{23}$$

where pc is the marginal cost in percent of the drug price.

Using this formula, we can estimate the cutoff $\bar{c}_{R\&D}(q_i)$ by assuming that bargaining power equals $\alpha = 50\%$ and percent marginal cost equals $pc = 10\%$ (see Leighl et al. (2021)).

However, we need a further adjustment. Drug prices outside the U.S. are lower than U.S. prices (see e.g. Goldstein et al. (2017)). Furthermore, targeted therapies are not as broadly available in many foreign countries as they are in the U.S., due to stricter regulatory requirements, and/or high costs. To jointly capture these price and market penetration differences, we leverage the fact that the U.S. accounts for about 45% of the total global spending in cancer drugs (see IQVIA, 2024, exhibit 53.) This implies that $45\% = \frac{P_i \times N^{US}}{P_i \times x \times N^W}$ where N^{US} and N^W are the U.S. and world population of lung cancer patients respectively, and x is the adjustment factor capturing price and market penetration differences. Since the

[14] 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
New cases per year	$\overline{N}$	2,500,000
Average Percent of Tested Patients	adj	0.5
Actual new cases per year	N	1,250,000
Eligibility for RCT		10%
Adjustment to N for RoW	x	0.196
Pharma's discount rate	R	10.5%
Length of clinical trial	τ	4 years
Length of Phase I	ρ	8 years
Pharma bargaining power	α	0.5
Expected standardized improvement	$E[\kappa]$	0.35
Probability $\kappa < 0$		0.3
Annual survival rate for T	$\overline{s}$	0.8
Annual survival rate for other causes	s_{other}	0.95
Corporate tax rate	t_c	21%
Orphan Drug tax credit	tc_{OD}	25%
Annual cost of RTC per patient	c_{RTC}	$\$100,000$
Success probability of Phase I	p_0	16.8%

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 take into account the lower prices and lower penetration by defining an effective world population $N = 0.196 \times N^W$ when we calculate profits or break-even prices.[15]

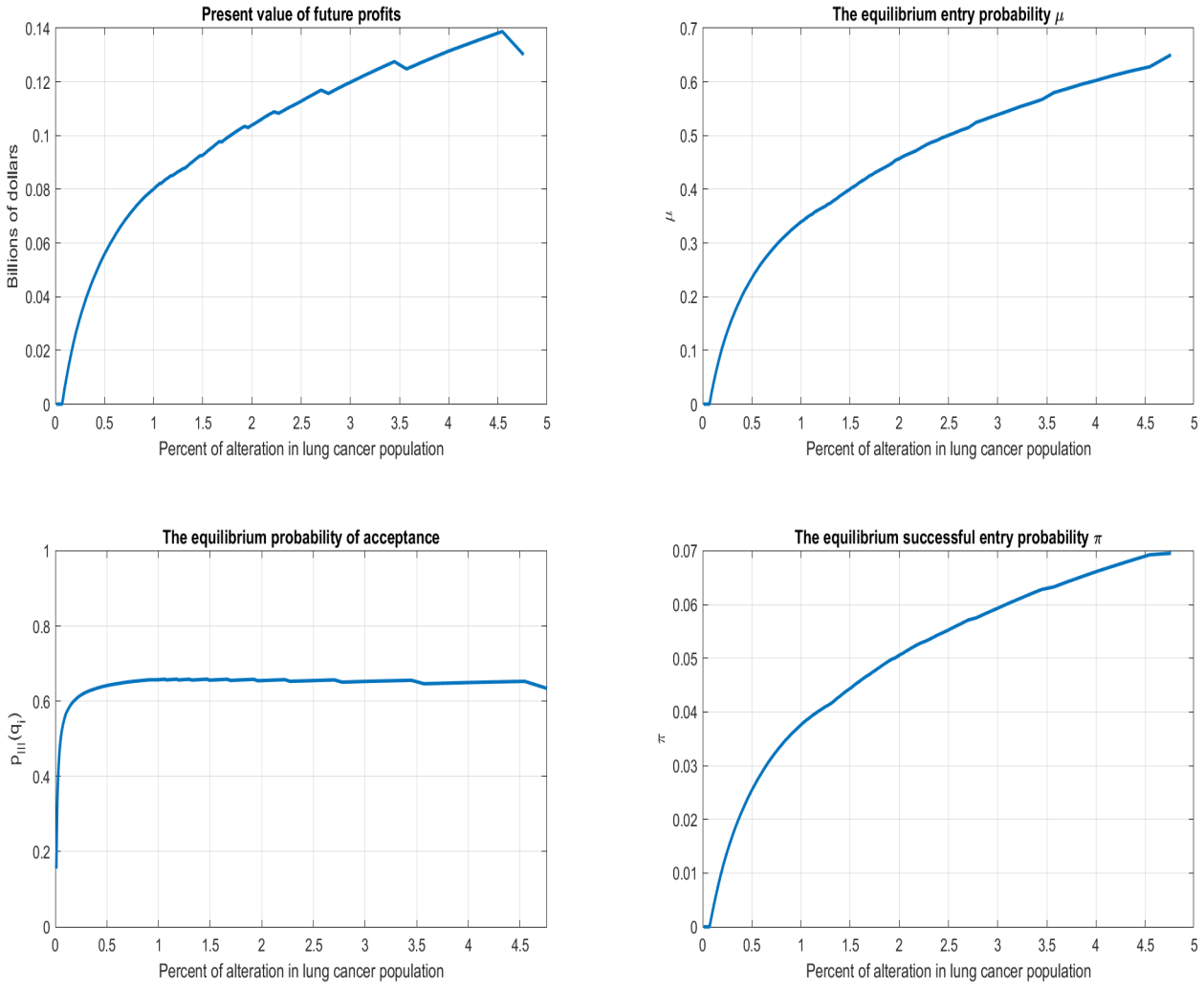
5.2. Calibration Results

The top-left panel of Figure 9 plots the present value of investing in the development of a treatment for a specific alteration, with R&D cost equal to the expectation $E[c_{R\&D}] = 0.5(c_L + \bar{c}_{R\&D})$. The horizontal axis reports the percentage of lung cancer patients with the given alteration, and the figure demonstrates that the present value converges to zero as the percentage of lung cancer patients affected by the alteration decreases. The present value hits zero at 0.07% of incidence. That is, alterations that affect less than 0.07% of lung cancer populations are not financially viable. This result already takes into account both the lower probability that a potential competitors may arrive due to lower incidence, as well as the higher expected drug prices for the same as shown in Figure 6.

The non-linearity that is evident in the top-left panel is due to the general equilibrium

[15]We could equivalently define an effective world price, but below we use the observed U.S. prices to estimate certain quantities, and such adjustments make the calculations non-transparent. An effective world population is easier to work with.

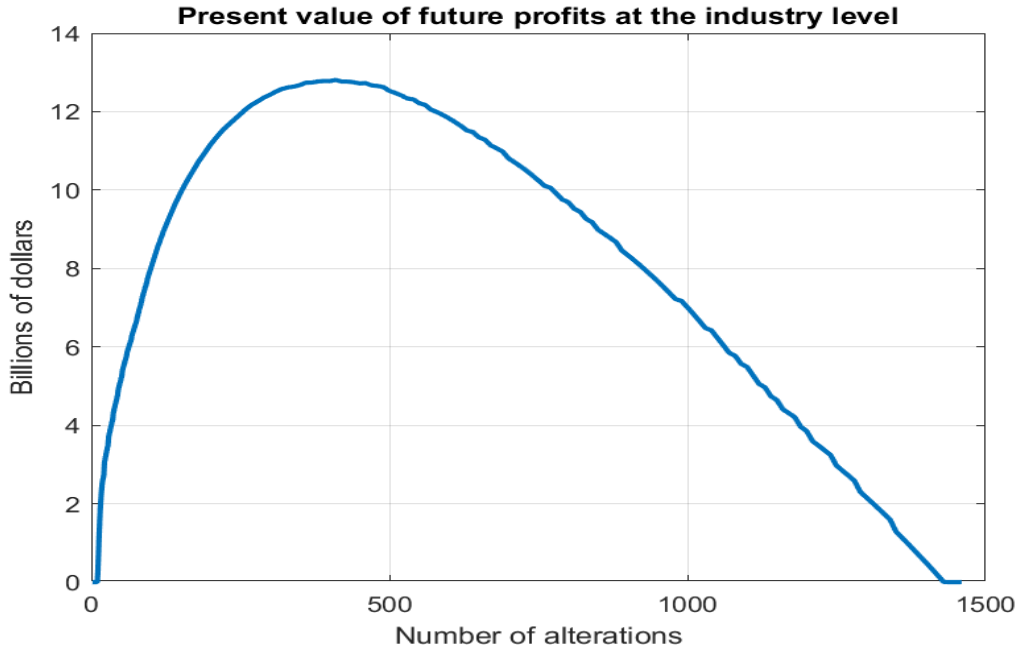
Figure 9: **Equilibrium Present value of Future Profits**



effect that rarer alterations imply a lower entry probability μ, as shown in the top-right panel. The lower entry probability implies a lower probability π that a new entrant will steal the market, as shown in the bottom-right panel. A lower probability of entry in turn yields an increase in the discount factors $A_{III}(\pi)$. The increase in discount factor increases the present value of future profits, everything else equal. Notwithstanding all these equilibrium counter effects, the present value of future profits declines to zero as the alteration becomes rarer, due to the lower number of patients available to sell the drug to, but also due to the lower probability of acceptance of the new drug, as shown in the bottom-right panel of the figure.

Cross-sectionally, these results also imply that we should observe more research intensity for alterations that affect a higher percent of the population. A higher μ for more common

Figure 10: **Equilibrium Present Value of Aggregate Industry Future Profits**



alterations implies that there are more firms doing experimentation for alterations that are more common. This is consistent with the data shown in Figure 5.

At the industry level, Figure 10 plots the present value of future profits against the number of discovered alterations n (see formula (22)). The plot shows that in aggregate, the present value initially increases as the number of alterations increases, due to higher price $P(n)$ and lower probability of entry π, which in turn increase the discount factor $A_{III}(\pi)$ in equilibrium. Eventually, however, the higher aggregate R&D costs $0.5(c_L + \bar{c}_{R\&D}) \times n$ and lower probability of acceptance make the aggregate profits at the industry level for the whole lung cancer diseases decrease all the way to zero. This occurs for a number of alterations equal to $n = 1440$, which corresponds to .07% of lung cancer population (see Figure 9.)

5.3. Calibration Results: The Social Costs of Personalized Medicine

We now compute social costs as in Section 3.12., assuming that the social planner aims to break even and cover the entire world population. 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 8). We consider three different levels of $R\&D$ costs, spanning from \$50 million to \$500 million. The latter figure is in line with the current ex-post costs of R&D as per Table 1.

More specifically, Panel A of Table 3 shows that as the number of discovered alterations increases, with each affecting a smaller number of patients, the break-even prices increase substantially. For instance, for $n = 1000$, which implies that each alteration affects 0.1% of the lung cancer population, the minimum break-even price for R&D costs per drug of just $c_{R\&D} = \$50$ million is $P^{BE} = \$304,750$ per year, per patient. Assuming more realistic R&D costs of \$500 million per drug, the break-even price rises to $P^{BE} = \$2.5$ million per year, per patient. When $n = 2,000$, for an incidence of 0.05% in the lung cancer population, the break-even prices equal $P^{BE} = \$588,784$ and \$5.4 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 $n = 2,000$, which entails each alteration has an incidence of 0.05%. Recall that this is realistic, in view of the data in Figure 1. In this case, societal costs would be \$719 billion per year when $c_{R\&D} = \$50$ million, and \$6.6 trillion per year when $c_{R\&D} = \$500$ million. And this only for lung cancer.

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).[16] 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 reveals that annual break-even costs of developing targeted therapies are higher than these values once the number of alterations n is over 2000, i.e. for an incidence of 0.05%.

6. Existing Solutions

The concern for the lack of R&D for the treatment of rare diseases is not new. (What is new is the realization that targeted therapies are progressively turning many diseases into collections of rare diseases.) For instance, Pizzamiglio et al. (2022) lists different trial designs that have been put forward to date for rare diseases. The designs considered therein include: (1) Cross-over, in which participants receive a random sequence of different treatments, each followed by a "wash-out" period, so that participants are effectively their own controls; (2) Delayed start, in which initially only a placebo-controlled phase is carried out, and treatment occurs during a second phase; (3) Randomized withdrawal, in which all participants first

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

Table 3: **Break-even Prices and Global Societal Costs of Targeted Therapies**

Panel A: Breakeven Annual Cost per Patient				
Number n	Percent in	Cost of pre-Phase III R&D		
of alterations	Population	50 million	100 million	500 million
10	10	40,197	54,336	167,447
20	5	50,212	70,427	232,149
50	2	69,760	102,020	360,098
100	1	94,810	143,301	531,230
200	0.5	127,451	202,737	805,019
500	0.2	196,274	337,508	1,467,378
1000	0.1	304,750	554,724	2,554,523
2000	0.05	588,784	1,122,385	5,391,194
5000	0.02	2,033,349	4,009,175	19,815,786
10000	0.01	8,563,760	17,061,712	85,045,333
Panel B: Breakeven Societal Cost (million of USD)				
Number n	Percent in	Cost of pre-Phase III R&D		
of alterations	Population	50 million	100 million	500 million
10	10	49,129	66,410	204,658
20	5	61,370	86,078	283,737
50	2	85,263	124,691	440,119
100	1	115,879	175,146	649,282
200	0.5	155,774	247,789	983,912
500	0.2	239,891	412,510	1,793,462
1000	0.1	372,472	677,997	3,122,195
2000	0.05	719,625	1,371,804	6,589,237
5000	0.02	2,485,204	4,900,103	24,219,294
10000	0.01	10,466,818	20,853,204	103,944,296

receive treatment, and then randomization is only carried out among those that initially respond; (4) Group sequential, in which the number of participants is not pre-set, but the RCT is monitored through pre-determined interim analysis which determines the possible need to terminate early; (5) Adaptive, in which the probability of randomization to one group shifts towards more promising treatments, based on partial results.

From the perspective of our model, these methodologies can help in that they decrease the total fixed R&D costs (e.g. through parallel testing), and the total number of patients M needed per alteration. However, Table 3 suggests that, even for low costs of $R\&D$, as the number of alterations n increases, the break-even price increases by several orders of magnitude. It seems highly unlikely that the savings resulting from the adoption of the methodologies just described would be able to offset this effect. Indeed, the quotations at the beginning of this article suggest that even regulatory agencies realize that these methodologies will not be sufficient for rare alterations. We need a new paradigm, discussed next.

7. A New Paradigm: The Personalized 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.

7.1. 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.

7.2. Competition in PDDPs: Economic Viability

We now analyze the economic viability of the PDDP. Specifically, we envision pharma companies developing competing PDDPs which may differ from each other due to differences in technologies, algorithms, datasets, and computing power. In a sense, it is as if each pharma company "becomes like Google" as at its core are algorithms and computing power.[17] While it would not be meaningful to compare a drug developed by one PDDP to another specific drug developed by another PDDP (because each drug is targeted to a different patient), we can envision comparing different PDDPs in the same way now pharma companies compare different drugs for the same alteration—by comparing their effectiveness for a sample of patients. Indeed, regulatory authorities such as the FDA and EMA would want to have approval power over PDDPs and therefore ask pharma companies to run randomized clinical trials to ensure that (a) PDDPs are better than the standard of care; and (b) any new PDDP improves over the PDDP currently in place. Crucially, the RCT is run over the entire lung cancer population, irrespective of which targetable alteration, if any, any single recruited patient has. Thus, the economic viability of a PDDP is independent of the rarity of the various alterations.

More concretely, the randomized clinical trial would be set up in the same format as in Section 3.. It is still convenient to divide a pharma company's decision to enter the PDDP's market in two phases, phase I and phase III. In phase I, the firm must acquire the necessary know-how—hire human capital, build or secure access to computing power, develop algorithms, and buy datasets. Let c_{tech} be the initial fixed costs of Phase I. Once the pharma company has acquired the necessary know-how, in phase III, it must run a RCT. As in Section 3., M patients are divided into a test group T and a control group C. The C group

[17] It is recent news that Google Deepmind platform help identify a novel cancer immunotherapy approach. See e.g. https://trial.medpath.com/news/47d1bfeceb1cbe0d/google-deepmind-s-ai-model-discovers-novel-cancer-immunotherapy-approach-through-cell-level-analysis

will receive the best existing treatment, or the currently best PDDP; for short, we refer to this as the "standard of care." The T group undergoes steps A through E in Section 7.1. of the personalized drug discovery process (including, if necessary, the standard of care as a fall-back option). The patients in the two groups are then followed for τ periods. Crucially, we do not differentiate across alterations. All patients with lung cancer are considered as just one group. All other steps in the RCT are otherwise the same, and after the observation period of τ periods the usual calculations are performed.

Let Q_T be the probability that the new PDDP detects a new alteration and develops a treatment for it, or develops a new drug for a known alteration. s_T^{new} denotes the survival probability of the new drug. As in previous sections, we define κ as the standardized improvement of T versus C, and we assume $\kappa \sim U[\kappa_L, \kappa_H]$ as in previous sections. In fact, we make the same distributional assumptions of T versus C to focus on the order of magnitudes of prices.

The present value of future profits at the beginning of phase III is then

$$PV_{III}^{PDDP} = p_{III}(M) \left(Q_T \, A_{III} \, (P^{New} - c^{New}) - C_{PDDP} \right) N - c_{RCT} \, M \tag{24}$$

where $p_{III}(M)$ is the probability of approval, given a sample of size M; A_{III} is the discount factor given by equation (4); and c_{PDDP} is the (variable) cost of running the PDDP, once the fixed costs have all been paid up. We make the pessimistic assumption that the pharma company makes profits on a patient only if it detects a new alteration or develops a new drug (which occurs with probability Q_T): in case it does not, the patient receives the standard of care, which we assume is delivered by a different pharma company. This of course minimizes the possible profits of the pharma company.

The decision to enter the market depends on the cost of acquiring the know-how, which we assume to be heterogeneous across pharma companies. We then have

$$PV_I^{PDDP} = A_I P_{III}^{PDDP} - c_{tech} \tag{25}$$

where $A_I = (1+r)^{\rho}$ is the time discount factor to acquire the necessary know-how. We assume that the pharma company can always procure the necessary hardware and software, access datasets and so on: the only uncertainty is whether their new PDDP will be better than the standard of care. As before, we assume that there is a a probability μ that a potential entrant arrives in the PDDP market, with heterogeneous $c_{tech} \sim U[c_{tech}^L, c_{tech}^H]$.

We continue to assume that the price of drugs from the PDDP, P^{new} above, is determined by bargaining between the pharma company and a health insurance company. Specifically, recall that, in every period, the insurance company sets up a set of maximum prices $\overline{P}_i$, one for each alteration i, designed so as to break even. With PDDP, we now also have one more

maximum price $\overline{P}^{new}$. With PDDP, the break-even condition is

$$\frac{C}{q} = Q_T \overline{P}^{new} + (1 - Q_T) \sum_{i=1}^{n} q_i \overline{P}_i \tag{26}$$

This formula takes into account that if PDDP fails to find a new drug, there is a fall-back strategy to the standard of care.[18] As in section 3.9., these maximum prices are set up ahead of the bargaining with the pharma companies. Assuming again an $\alpha-$split of the bargaining between pharma company and health insurance company, the price of PDDP is given by

$$P^{new} = \alpha \, \overline{P}^{new} + (1 - \alpha) \, P^{BE}_{PDDP} \tag{27}$$

where P^{BE}_{PDDP} is the breakeven price from setting $PV_I^{PDDP} = 0$ in (25), explicitly given by

$$P^{BE}_{PDDP} = c^{new} + \frac{c_{tech} + A_I c_{RCT} M^*}{A_I p_{III}(M^*) Q_T A_{III} N} + \frac{c_{PDDP}}{Q_T A_{III}} \tag{28}$$

Important, neither the present value PV_I^{PDDP} equation (25) nor the break-even price P^{BE}_{PDDP} in (28) depend on the number of alterations n. That is, as n increases, nothing changes in these quantities. Moreover, we have the following proposition:

Proposition 10 *Everything else equal, the present value of future profits PV_I^{PDDP} in equation (25) increases, and the break-even price P^{BE}_{PDDP} in equation (28) declines, as the ability of PDDP of detecting a new alteration and find a new drug Q_T increases.*

The increase in PV_I^{PDDP} as Q_T increases entails that pharma companies have an incentive to keep exploring new systems to increase Q_T over their competitors. More specifically, given a probability μ of arrival of a potential entrants, the total probability of a competitor stealing the PDDP market is $\pi = p(M^*)\,\mu$. Given heterogeneous costs $c_{tech} \sim U[c^L_{tech}, c^H_{tech}]$, the probability μ itself equals the probability of arrival of a pharma company with $PV_I^{PDDP} > 0$, that is, whose cost c_{tech} is below the threshold

$$\overline{c}_{tech}(\pi) = A_I(\pi) P^{PDDP}_{III}(\pi) \tag{29}$$

where we emphasize that the right-hand-side depends on π itself through the discounts $A_I(\pi)$ and $A_{III}(\pi)$. As in section 3.10. there is a fixed point $\mu \to \pi \to \mu$

Proposition 11 *There is a fixed point such that $\mu(\pi) = \frac{\overline{c}_{tech}(pi) - c^L_{tech}}{c^H_{tech} - c^L_{tech}}$ and $\pi = p(M)\mu(\pi)$.*

In the next section, we calibrate the model to provide insights about the magnitudes.

[18] Eq. 26 is written for the case in which the standard of care consists of targeted therapies for each alteration. This facilitates the comparison with our prior analysis. The analysis for the case in which the standard of care is a previously approved PDDP is analogous, except that the summation in Eq. 26 is replaced by the maximum price for the previous PDDP.

7.3. Calibration

The model in the previous section involves several uncertain quantities—PDDP has not been implemented yet. To address this, we calculate values for a range of parameter values for some of the key quantities. First, we allow Q_T to range between .1 and .9. Second, we keep the quantities related to the phase–III RCT as in previous section: namely, $\tau = 4$ years, $c_{RCT} = \$100,000$ per patient, and the probability of approval is calculated assuming κ has $E[\kappa] = 0.35$ and $pr(\kappa < 0) = 0.3$, as before. Although speculative, these values provide some baseline for the prices P^{new} we are looking for, and allow a comparison with prices under the current regulatory regime. We also assume it takes $\rho = 2$ years to acquire the know-how, hardware, software, and data in Phase I. Finally, we assume $c_{PDDP} = \$1000$, and $c_{tech} \sim U[c^L_{tech}, c^H_{tech}]$ where $c^L_{tech} = \$500$ million and $c^H_{tech} = \$1.5$ billion. For reference, the cost of the Aurora super computer at the Argonne National Laboratory was \$500 million to build. The marginal cost c^{new} is set conservatively equal to the one used for the rarest alteration in Section 5., namely $c^{new} = 10\% P^{NTRK}$, where the latter is the price for Lorlatinib for NTRK alteration.

In order to calculate P^{new} we need to also calculate $\overline{P}^{new}$. Because $C/q = \sum_{i=1}^n \overline{P}_i\, q_i$ at the time the first PDDP undergoes an RCT, and the standard of care at that time is indeed to target individual alterations as in Section 3.9., we indeed have $\overline{P}^{new} = \sum_{i=1}^n \overline{P}_i\, q_i$. We can estimate the latter max prices $\overline{P}_i$ from observed prices, alteration by alteration, by matching $P_i = P_i^{data}$ in equation (17), and then inverting the same equation to recover $\overline{P}_i$.

Table 4 reports the prices P^{new} and the present value of future profits PV_I^{PDDP} under three assumptions about the survival rate s_{PDDP} and two assumptions about the bargaining power α (Panels A and B). Consider Panel A. First, we see that the PDDP prices range between \$275,139 and \$253,370 per year for $s_{PDDP} = 0.7$. These prices are slightly higher than the current EGFR prices ($\$228,000$), and about half of the prices of the most expensive drugs currently in commerce (see Figure 6). Importantly, for given effectiveness s_{PDDP}, the equilibrium price is lower the higher the probability Q_T of finding a new treatment. The reason is that a higher probability implies a lower break-even price for the pharma company, and thus less effective "bargaining power." Still, even with lower prices, the present value of future profits increases as Q_T increases: In fact, for each s_{PDDP}, it increases tenfold from $Q_T = .1$ to $Q_T = .9$. For instance, for $s_{PDDP} = 0.8$, the present value of future profits increases from \$1 billion for $Q_T = 0.1$ to \$11 billion for $Q_T = 0.9$. This is in line with proposition 10.

Similarly, as we move across survival rates s_{PDDP} we see that PDDP prices are rather similar, but the present value of future profits increase substantially. For instance, for $Q_T = 0.5$, PV_I^{PDDP} increases from \$5.32 billion to \$7.75 billion, nearly a 50% increase.

Table 4: **Forecasted PDDP Prices and Present Value of Future Profits**

	PDDP Prices			PV of Future Profits (Billions of USD)		
	Panel A: Bargaining Power $\alpha = 0.5$					
Q_T	s_{PDDP} =.7	s_{PDDP} =.8	s_{PDDP} =.9	s_{PDDP} =.7	s_{PDDP} =.8	s_{PDDP} =.9
0.1	275,139	272,565	269,555	0.96	1.15	1.44
0.2	262,894	262,109	261,142	2.05	2.43	3.02
0.3	258,813	258,623	258,338	3.14	3.71	4.60
0.4	256,772	256,881	256,936	4.23	5.00	6.17
0.5	255,547	255,835	256,095	5.32	6.28	7.75
0.6	254,731	255,138	255,534	6.41	7.56	9.32
0.7	254,148	254,640	255,134	7.50	8.85	10.90
0.8	253,711	254,267	254,833	8.59	10.13	12.47
0.9	253,370	253,976	254,599	9.68	11.41	14.05
	Panel A: Bargaining Power $\alpha = 0.9$					
Q_T	s_{PDDP} =.7	s_{PDDP} =.8	s_{PDDP} =.9	s_{PDDP} =.7	s_{PDDP} =.8	s_{PDDP} =.9
0.1	250,200	249,672	249,062	0.82	1.00	1.29
0.2	247,751	247,581	247,379	1.88	2.25	2.81
0.3	246,934	246,884	246,818	2.94	3.49	4.33
0.4	246,526	246,535	246,538	4.00	4.73	5.85
0.5	246,281	246,326	246,370	5.06	5.98	7.37
0.6	246,118	246,187	246,257	6.12	7.22	8.89
0.7	246,001	246,087	246,177	7.19	8.46	10.41
0.8	245,914	246,012	246,117	8.25	9.71	11.93
0.9	245,846	245,954	246,070	9.31	10.95	13.45

These results show that it is in the interest of the new "Googlefied" pharma companies to find increasingly better system that provide new personalized therapies and higher survival rates. Indeed, in all cases, the equilibrium fraction of firms that decide to enter μ equals one. This is in stark contrast with the current system, under which firms have an incentive to lower the investments in rare alterations, including avoiding to enter the market altogether.

8. 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.

As the quotes at the beginning of this article show, regulatory agencies such as the FDA in U.S. and the MHRA in U.K. are indeed acknowledging the limitations inherent in the current approval system and, consistently with our proposal, are considering "process-based" trials. Our paper provides a blueprint for the economic analysis of such an approval regime, especially as it compares to the current one.

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] 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.

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

[5] 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.

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

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

[8] 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.

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

[10] 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.

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

[12] 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.

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

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

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

[16] 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.

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

[18] 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.

[19] 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.

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

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

[22] 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).

[23] 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.

[24] Chiara Pizzamiglio et al. “Designing clinical trials for rare diseases: unique challenges and opportunities”. In: *Nature Reviews Methods Primers* (2022), pp. 1–2.

[25] 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.

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

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

[28] 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.

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

[30] JM Unger et al. “National Estimates of the Participation of Patients With Cancer in clinical Research Studies Based on Commission on Cancer Accreditation Data.” In: *J Clin Oncol.* 42.18 (2024), pp. 2139–2148.

[31] 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.

[32] 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.

[33] 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.

[34] 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:

$$Prob\left(X_T = x_T | X_T + X_C = t, H_0\right) = \frac{\begin{pmatrix} M/2 \\ x_T \end{pmatrix}\begin{pmatrix} M/2 \\ t - x_T \end{pmatrix}}{\begin{pmatrix} M \\ t \end{pmatrix}}. \tag{30}$$

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 (30) 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 $\kappa \sim Uniform[a, b]$ with a and b chosen to have

$E[\kappa] = 0.35$ and $Pr(\kappa < 0) = 0.3$. We set $p_C = 0.35$ and then for each simulated $\tilde{\kappa}$ we invert the function $\tilde{\kappa} = (\tilde{p}_T - p_C)/\sqrt{p_C(1-p_C) + \tilde{p}_T(1-\tilde{p}_T)}$ to obtain simulated samples of $\tilde{p}_T$. 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 (30).[19]

- 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

$$\Pr\left(Approval|K,n\right) = 1 - \sum_{M=2}^{K} Pr(M|K)Pr(Approval|M,n) \tag{31}$$

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

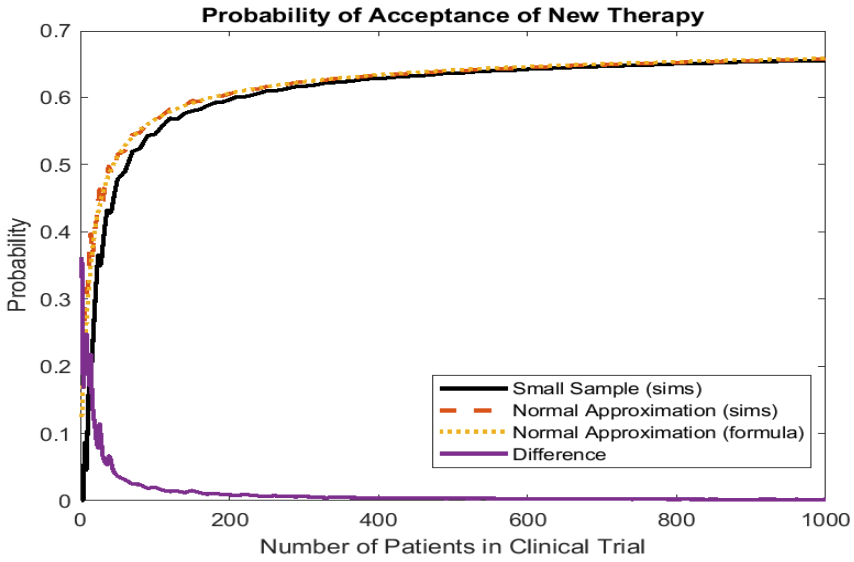
- We then repeat the whole procedure above for $n = 2$ to $n = 10000$.

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.

[19] 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 3.7., 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

$$g(M) \equiv \begin{cases} \Phi\left(z(\alpha) - \kappa\sqrt{\frac{M}{2}}\right) & M \geq 0 \\ 1 & M < 0. \end{cases} \tag{32}$$

B1. Proof of Proposition 1

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

$$PV(K,n) = (1 - \mathrm{E}[g(M)]) \cdot C - K \cdot c_{RCT} \tag{33}$$

where

$$C = A(P_{TT} - c)N. \tag{34}$$

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

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

Differentiating wrto K yields

$$\begin{aligned}
&\int \frac{-\frac{1}{2}K^{-\frac{3}{2}}}{\sqrt{2\pi q}} e^{-\frac{1}{2Kq}(M-Kq)^2} g(M)\, dM \\
&+ \int \frac{1}{\sqrt{2\pi Kq}} e^{-\frac{1}{2Kq}(M-Kq)^2} \left\{ -\frac{-K^{-2}}{2q}(M-Kq)^2 - \frac{1}{2Kq} 2(M-Kq)(-q) \right\} \\
&\quad \cdot g(M)\, dM \\
&= \int \left(-\frac{1}{2K}\right) \phi_{Kq,Kq}(M) g(M)\, dM \qquad (35) \\
&+ \int \left(\frac{1}{2K^2 q}(M-Kq)^2\right) \phi_{Kq,Kq}(M) g(M)\, dM \qquad (36) \\
&+ \int \left(\frac{1}{K}(M-Kq)\right) \phi_{Kq,Kq}(M) g(M)\, dM. \qquad (37)
\end{aligned}$$

Adding up the terms in parentheses in Eqs. (35)–(37) yields

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

Moreover,

$$\begin{aligned} &- Kq + (M - Kq)^2 + 2Kq(M - Kq) \\ =& - Kq + M^2 - 2MKq + K^2q^2 + 2MKq - 2K^2q^2 \\ =& - Kq + M^2 - K^2q^2 \\ =& - \mathrm{Var}[M] + M^2 - \mathrm{E}[M]^2 = M^2 - \mathrm{E}[M^2]. \end{aligned} \tag{38}$$

Therefore,

$$\frac{\partial E[g(M)]}{\partial K} = \frac{\mathrm{E}\left[(M^2 - \mathrm{E}[M^2]) \cdot g(M)\right]}{2K^2q} < 0. \tag{39}$$

We now provide a complete argument for the last inequality. Note first that because $g(M)$ has a jump discontinuity at $M = 0$, we cannot simply apply Stein's Lemma ($\mathrm{E}[(M-m)f(M)] = v\mathrm{E}[f'(M)]$ for $M \sim \mathcal{N}(m, v)$) with $f = g$. Instead, we proceed by first establishing the result for smooth functions and then using a bounding argument.

Step 1: The Smooth Case. Let $h : \mathbb{R} \to [0, 1]$ be a differentiable, weakly decreasing function. We expand the centered quadratic $M^2 - \mathrm{E}[M^2]$ as $(M - m)^2 - v + 2m(M - m)$. Applying Stein's Lemma to the terms separately:

$$\begin{aligned} \mathrm{E}[2m(M - m)h(M)] &= 2mv\mathrm{E}[h'(M)] \\ \mathrm{E}[((M - m)^2 - v)h(M)] &= \mathrm{E}[(M - m)\{(M - m)h(M)\}] - v\mathrm{E}[h(M)] \\ &= v\mathrm{E}[h(M) + (M - m)h'(M)] - v\mathrm{E}[h(M)] = v\mathrm{E}[(M - m)h'(M)]. \end{aligned}$$

Summing these evaluated terms yields $\mathrm{E}[(M^2 - \mathrm{E}[M^2])h(M)] = v\mathrm{E}[(M + m)h'(M)]$. Since h is weakly decreasing, $h'(M) \leq 0$ everywhere, and $M + m > 0$ for all $M \geq 0$, the integral is strictly negative.

Step 2: Handling the Discontinuity. For the original $g(M)$, we fix a small $\epsilon > 0$ such that $M^2 - \mathrm{E}[M^2] < 0$ for $M \in [-\epsilon, \epsilon]$. We construct a smooth, differentiable bridging function $h(M)$ such that $h(M) = g(M)$ for $|M| > \epsilon$ and $h(M) \leq g(M)$ for $M \in [-\epsilon, \epsilon]$. Since $M^2 - \mathrm{E}[M^2]$ is negative on this interval, multiplying by the lower bound $h(M)$ reverses the inequality: $(M^2 - \mathrm{E}[M^2])g(M) \leq (M^2 - \mathrm{E}[M^2])h(M)$. Taking expectations preserves the bound: $\mathrm{E}[(M^2 - \mathrm{E}[M^2])g(M)] \leq \mathrm{E}[(M^2 - \mathrm{E}[M^2])h(M)] < 0$.

Step 3: Constructing the Bridging Function. We define $h(M)$ on $[-\epsilon, \epsilon]$ by its derivative $h'(M) = \frac{1}{2\epsilon}[-Ct(1 - t)^k + Dt^m]$ where $t = (M + \epsilon)/(2\epsilon)$. We choose $D = 2\epsilon g'(\epsilon)$ to match the slope at $M = \epsilon$ and C to match the total drop $\Delta = g(-\epsilon) - g(\epsilon)$. By choosing k sufficiently large, $h(M)$ drops rapidly immediately after $M = -\epsilon$, ensuring $h(M) \leq g(M)$

and tracing the lower bound of the box just above $g(\epsilon)$. This successfully handles the jump discontinuity at zero. *Q.E.D.*

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

B2. Proof of Proposition 2

Let $q < q'$ and denote by K and K' the optimal choices of sizes for q and q', respectively. By Proposition 1,

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

furthermore, since K' is optimal for q, $PV_{III}(K',q') \geq PV_{III}(K,q')$. Thus, as claimed, $PV_{III}(K,q) < PV_{III}(K',q')$. *Q.E.D.*

B3. Proof of Proposition 3

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

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

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

B4. Proof of Proposition 4

The claim follows from Proposition 1 and equation 16. The last part of the claim follows from $P \leq \overline{P}$ and the latter being finite. Q.E.D.

B5. Proof of Proposition 5

Note first that a higher $\mu(q_i)$ implies a higher $\pi(q_i)$ from (19), and thus lower discount factors $A_I(\pi)$ and $A_{III}(\pi)$. This in turn implies $PV_I(q_i)$ is lower for higher $\mu(q_i)$. This implies that the cut-off R&D cost $\overline{c}_{R\&D}$ for which $PV_I(q_i) = 0$ is lower. Hence, given the uniformly distributed costs $c_{R\&D}$, the entry probability is $\mu(q_i) = \frac{\overline{c}_{R\&D}-c_L}{c_H-c_L}$ is lower. It follows that $\mu : [0,1] \to [0,1]$ is a contraction and from the contraction mapping theorem, there is a unique fixed point. Q.E.D.

B6. Proof of Proposition 6

The claim follows from the fact that for given set of parameters, there exists q_i sufficiently small for which $PV_I(q_i) < 0$. From Proposition 5 the only fixed point must have $\mu(q_i) = 0$. Q.E.D.

B7. Proof of Proposition 7

The proof is identical to the proof of Proposition 3. Q.E.D.

B8. Proof of Proposition 8

Setting $E[PV_I] = 0$ we obtain

$$P^{BE} = c_p + \frac{p_I\, A_I(n)\, c_{RCT}\, K + c_{R\&D} n}{p_I\, A_I(n)\, p(K^*(n), n)\, A_{III}(n)\, N}. \tag{40}$$

As n increases, the equilibrium probability of a new entry declines and thus the discounts $A_I(n)$ and $A_{III}(n)$ decrease. Still, as $n \to \infty$ at most $\pi \to 0$ which still leads to finite A_I and A_{III}. Since $q \to 0$ implies $p(K^*(n), n) \to 0$, the claim follows from the expression of P^{BE} in (40). Q.E.D.

B9. Proof of Proposition 9

First, for given $c_{R\&D}$, the price is

$$\begin{aligned}
P_i &= (1-\alpha)\, P_i^{BE} + \alpha \overline{P}_i \\
P_i &= (1-\alpha)\, c_i + (1-\alpha)\, \frac{p_I A_I (1-t_{CR}) c_{RCT} K q_i + c_{R\&D}}{p_I A_I p\,(K, q_i)\, A_{III} N q_i} + \alpha \overline{P}_i \\
P_i &= c_i + (1-\alpha)\, \frac{p_I A_I (1-t_{CR}) c_{RCT} K q_i + c_{R\&D}}{p_I A_I p\,(K, q_i)\, A_{III} N q_i} + \alpha \left(\overline{P}_i - c_i\right)
\end{aligned}$$

For given $c_{R\&D}$, the present value formula is

$$PV_I = p_I A_I p\,(K^*, q_i)\, A_{III}\, (P_i - c_i)\, N q_i - p_I A_I\, (1 - t_{CR})\, c_{RCT} K^* q_i - c_{R\&D}$$

or, substituting $P_i - c_i$ from previous formula

$$\begin{aligned}
PV_I &= p_I A_I p\,(K^*, q_i)\, A_{III} \left((1-\alpha)\, \frac{p_I A_I (1-t_{CR}) c_{RCT} K q_i + c_{R\&D}}{p_I A_I p\,(K, q_i)\, A_{III} N q_i} + \alpha \left(\overline{P}_i - c_i\right) \right) N q_i \\
&\quad - p_I A_I\, (1 - t_{CR})\, c_{RCT} K^* q_i - c_{R\&D}
\end{aligned}$$

Tedious algebra gives

$$\begin{aligned} PV_I &= -\alpha c_{R\&D} + p_I A_I p\left(K^*, q_i\right) A_{III} \alpha \left(\overline{P}_i - c_i\right) N q_i \\ &\quad -\alpha p_I A_I \left(1 - t_{CR}\right) c_{RCT} K^* q_i \end{aligned}$$

Therefore, we have $PV_I = 0$ if and only if the cutoff $\overline{c_{R\&D}}$ is given by:

$$\overline{c_{R\&D}} = \left(p_I A_I p\left(K, q_i\right) A_{III} \left(\overline{P}_i - c_i\right) N - p_I A_I \left(1 - t_{CR}\right) c_{RCT} K\right) q_i \tag{41}$$

For a given value of $\overline{c_{R\&D}}$, we can now calculate the expected $c_{R\&D}$ as

$$E\left[c_{R\&D}\right] = \int_{c_L}^{\overline{c_{R\&D}}} \frac{x}{\overline{c_{R\&D}} - c_L} dx = \frac{1}{2} \frac{\overline{c_{R\&D}}^2 - c_L^2}{\overline{c_{R\&D}} - c_L} = \frac{1}{2}\left(\overline{c_{R\&D}} + c_L\right)$$

If the observed price $\widehat{P}_i$ is determined by the firm with expected cost $E[c_{R\&D}]$ (the average firm), we have that the expected bargaining price would be given by

$$E\left[P_i\right] = c_i\left(1 - \alpha\right) \frac{p_I A_I (1 - t_{CR}) c_{RCT} K q_i + E\left[c_{R\&D}\right]}{p_I A_I p\left(K, q_i\right) A_{III} N q_i} + \alpha\left(\overline{P}_i - c_i\right)$$

or, defining $pc_i = c_i / P_i$ and substituting for $E[c_{R\&D}]$,

$$E\left[P_i\right]\left(1 - pc_i\right) = \left(1 - \alpha\right) \frac{p_I A_I (1 - t_{CR}) c_{RCT} K q_i + \frac{1}{2}\left(\overline{c_{R\&D}} + c_L\right)}{p_I A_I p\left(K, q_i\right) A_{III} N q_i} + \alpha\left(\overline{P}_i - c_i\right)$$

Imposing then

$$\widehat{P}_i = E\left[P_i\right]$$

we find that for given $\overline{c_{R\&D}}$

$$\left(\overline{P}_i - c_i\right) = \widehat{P}_i \frac{\left(1 - pc_i\right)}{\alpha} - \frac{\left(1 - \alpha\right)}{\alpha} \frac{p_I A_I (1 - t_{CR}) c_{RCT} K q_i + \frac{1}{2}\left(\overline{c_{R\&D}} + c_L\right)}{p_I A_I p\left(K, q_i\right) A_{III} N q_i}$$

It follows from (41) that

$$\overline{c_{R\&D}} = \left(\begin{array}{c} p_I A_I p\left(K, q_i\right) A_{III} \left(\widehat{P}_i \frac{(1 - pc_i)}{\alpha} - \frac{(1-\alpha)}{\alpha} \frac{p_I A_I (1 - t_{CR}) c_{RCT} K q_i + \frac{1}{2}(\overline{c_{R\&D}} + c_L)}{p_I A_I p(K, q_i) A_{III} N q_i}\right) N \\ -p_I A_I \left(1 - t_{CR}\right) c_{RCT} K \end{array} \right) q_i$$

We can now solve for $\overline{c_{R\&D}}$. After some tedious algebra, we obtain

$$\overline{c_{R\&D}} = \frac{2}{1 + \alpha} p_I A_I \left(p\left(K, q_i\right) A_{III} \widehat{P}_i \left(1 - pc_i\right) N - (1 - t_{CR}) c_{RCT} K\right) q_i - \frac{(1 - \alpha)}{1 + \alpha} c_L$$

Q.E.D

B10. Proof of Proposition 10

The claim follows immediately from the expressions for PV_I^{PDDP} and the break-even price P_{PDDP}^{BE}. Q.E.D.

B11. Proof of Proposition 11

The argument is identical to the one of Proposition 5. Q.E.D.