CHARACTERIZING THE DRUG DEVELOPMENT PIPELINE FOR PRECISION MEDICINES

Amitabh Chandra[†] Craig Garthwaite[‡] Ariel Dora Stern[§]

Abstract

Precision medicines – therapies that rely on genetic, epigenetic, and protein biomarkers – create a better match between individuals with specific disease subtypes and medications that are more effective for those patients. These treatments are expected to be both more effective and more expensive than conventional therapies, implying that their introduction is likely to have a meaningful effect on health care spending. Using a comprehensive database of over 140,000 global clinical trials, we describe the drug development pipeline for precision medicines by characterizing drug development efforts over the past two decades. We identify clinical trials for *potential precision medicines* (PPMs) as those that use one or more relevant biomarkers. We then further segment trials based on the nature of the biomarker(s) used and other trial features with economic implications. Since cancers represent a set of diseases in which precision therapies are already successfully used, and since cancer applications of precision medicine are expected to grow rapidly over the coming years, we separately characterize cancer PPMs. We also summarize the role of National Institutes of Health (NIH) in supporting the existing pipeline of precision medicines, by asking what share of pipeline precision medicines rely on research supported by NIH grants. Finally, we consider the types of firms pursuing R&D in precision medicines, considering how PPM R&D activities have evolved over recent years.

Acknowledgements: We are grateful to Brian Alexander, Can Huang, Mark Trusheim, and participants in workshops for the NBER Project on Economic Dimensions of Personalized and Precision Medicine, the NBER-AIEA Conference, and the Bates White Life Sciences Symposium, for helpful comments. Ben Berger, Holly Breuer, Andrew Marder, Caroline Marra, and Alice Ndikumana provided excellent research and programming assistance.

† amitabh_chandra@harvard.edu, Harvard Kennedy School and NBER
‡ c-garthwaite@kellogg.northwestern.edu, Northwestern University Kellogg School of Management and NBER
§ astern@hbs.edu, Harvard Business School

1. INTRODUCTION

While lacking a universally agreed upon definition, Precision Medicine is broadly defined as an approach to disease treatment and prevention that takes into account individual variability in environment, lifestyle, and genes for each person.[1] The concept of targeted interventions has a long history across the practice of medicine, however, recent technological advancements have made it increasingly possible to tailor the development and utilization of medical technologies. This possibility has attracted broad interest from the medical and broader scientific communities. For example, in early 2015, the White House announced a "bold new research effort to revolutionize how we improve health and treat disease," and launched a Precision Medicine Initiative with a $215 million investment in 2016.[2] Other countries such as France and China have also announced major public investments ranging from the equivalent of several hundreds of millions of U.S dollars to several billion over coming years. Major investments to advance precision medicine have also been announced by a number of U.S. research institutions such as Harvard University and the University of California San Francisco.[3]

Below, we consider a subset of the broad set of practices encompassed by "precision medicine" and focus specifically on the clinical development of precision *medicines*, i.e. those new therapies focused on biomarker-defined patient subgroups. Precision medicines, and in particular, therapies that rely on genetic, epigenetic, and protein biomarkers, can help patients by using identifiable biological features (biomarkers) to define disease subtypes. The technology to rapidly and accurately sequence genes has increasingly facilitated understand the "-omic" (genomic and proteomic) characteristics of disease in recent years. This, in turn, has broadened the scope for drug development focusing on targeted therapies for newly-identifiable

[1] https://www.nih.gov/research-training/allofus-research-program
[2] https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative
[3] http://solidarites-sante.gouv.fr/IMG/pdf/genomic_medicine_france_2025.pdf
https://www.genomeweb.com/clinical-translational/france-plans-invest-670m-genomics-personalized-medicine
https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative
http://www.nature.com/news/china-embraces-precision-medicine-on-a-massive-scale-1.19108
http://www.hbs.edu/news/releases/Pages/kraft-family-foundation-establishes-endowment.aspx
https://www.ucsf.edu/news/2015/08/131341/new-center-will-advance-life-saving-genome-based-diagnostic-tools

sub-groups of patients. Indeed, the public efforts noted above lag private endeavors in this area: the pharmaceutical industry[4] has already commercialized almost 150 drugs with pharmacogenomic information in their label, according to the U.S. Food and Drug Administration (FDA).[5]

We focus on precision medicines because, in theory, they allow for a more effective match between individuals with specific disease sub-types and medications that are more effective for those sub-types. While the science underlying these medicines is broadly interesting and is the subject of a growing body of research, this ability to more precisely match patients and medications based on likely efficacy also fundamentally changes many of the *economic* incentives that pharmaceutical manufacturers face in the drug development process.

Perhaps most importantly, the ability to develop more targeted products may influence the decision process for which drugs to bring to market. These decisions will then subsequently be reflected in the equilibrium prices and availability of new pharmaceutical products. For example, almost by definition, precision medicines tend to target smaller patient populations than more traditional medicines. This may mean that manufacturers will shift their attention to the subset of products able to command high(er) prices – and thus are more likely to justify the fixed costs of developing the medication. These higher priced products are likely to include those products with large clinical benefits in relatively small patient populations. In addition, since these drugs are more efficacious within a smaller patient population, the marginal customer is expected to have a greater willingness to pay, allowing for higher profit maximizing prices on the part of manufacturers. These two factors together provide an economic rationale for the broadly higher prices observed for precision medicines.

Economic incentives could also, all else held equal, result in some products no longer being brought to market because manufacturers no longer believe they can reasonably expect to recuperate their research and development (R&D) expenditures from relatively small target patient populations. Potentially counteracting this effect, the ability to create identifiable subgroups of patients based on their willingness to pay increases the scope for future price-discrimination by manufacturers, who could, in theory, more easily charge higher prices for high-value indications and lower prices for indications or patients where therapies

[4] Throughout the chapter, reference to the pharmaceutical industry and pharmaceutical manufacturers refers to all firms developing drugs to treat medical conditions, including pharmaceutical and biotechnology firms

[5] http://www.personalizedmedicinecoalition.org/Userfiles/PMC-Corporate/file/pmc_personalized_medicine_by_the_numbers.pdf

will work less well (Chandra and Garthwaite, 2017). This would (weakly) increase the profits from any particular product and would, in turn, (weakly) increase the subset of early-stage products that pharmaceutical manufacturers would consider as candidates for commercialization. In addition, greater potential therapeutic benefit may result in smaller, shorter clinical trials because fewer patients would be needed and shorter periods of time will be sufficient for demonstrating statistically significant improvements in outcomes. Smaller and/or faster trials would both decrease the costs of bringing a drug to market and could increase the drug's effective patent length,[6] increasing the set of pipeline drugs considered as potentially worthwhile R&D investments.

Despite the potential for precision medicines to both reduce some of the costs of drug development and also increase the patient benefits created by new products, markets for some medicines may still be so small that private firms will lack the necessary incentives for bringing therapies to market. This would create a potential role for government funding of research in these areas from sources such as the National Institutes of Health (NIH).

Finally, the emergence of a new technology could create opportunities for additional specialization of firms into different stages of the development process. This could, for example, lead to early-stage drug discovery being increasingly pursued by a subset of highly specialized firms. More broadly, it is possible that the emergence of precision medicines will shift the division of labor between small biotechnology companies and large pharmaceutical companies across different stages of the R&D process.

To help understand this collection of potential economic implications of precision medicines, we aim to provide a detailed characterization of the existing drug development efforts in this area. We begin at a broad level by examining the aggregate development of potential precision medicines (PPMs), those pipeline drugs whose clinical trials have signature features of precision medicine R&D. We identify and report on clinical trials for such medicines by therapeutic area and over time. Since cancers represent a set of diseases in which precision therapies are already successfully used, and since cancer applications of precision medicine are expected to grow rapidly over the coming years, we separately characterize cancer PPMs.

[6] Patent life for a drug in the U.S. is generally 20 years from the date the application is filed and manufacturers can file a patent application any time before or during a drug's development process. Therefore, the time that a drug spends in clinical trials (i.e. before the drug can be marketed) are typically counted against the 20-year patent life. Marketing exclusivity is different from patent life and is granted by the FDA upon drug approval. Exclusivity typically lasts for 5 years, though there are extensions to exclusivity for certain cases, such as orphan drugs and pediatric indications. (https://www.fda.gov/downloads/drugs/developmentapprovalprocess/smallbusinessassistance/ucm447307.pdf)

Understanding the nature of these innovations provides first order information on the wide-ranging health care spending implications of these emerging medications.

We then examine other aspects of clinical trials that provide additional insight into the economic mechanisms of drug development that are shaping the nature of innovation in this area. We begin by considering the characteristics (e.g. geography, indication, sponsorship) of clinical research between PPM vs. non-PPM trials. We then summarize the role of the NIH in supporting the existing pipeline of precision medicines, by asking what share of PPM clinical trials cite the support of NIH grants. Finally, we consider the types of firms pursuing clinical trials in PPMs, considering how PPM R&D activities has evolved over recent years.

2. Precision Medicines and the Drug Development Process

As discussed above, we focus on the development of precision medicines – those products that use biomarkers to target particular subgroups of patients. To better understand how these products are defined and developed, we begin by providing some background information on the science of biomarkers and their use by various economic actors in the drug development process.

Precision medicines and biomarkers

The FDA defines a "biomarker" as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or biological responses to a therapeutic intervention."[7] A familiar example can be seen in the common medical practice of using glycated hemoglobin (HbA1c), an indicator of average blood glucose levels over time, as a measure of the effectiveness of a therapeutic agent in controlling diabetes. In this example, the biomarker (which indicates therapeutic efficacy) is HbA1c. However, biomarkers can also be used to carve out patient subtypes of diseases because a treatment may work differently in patients who vary in their biomarker subtypes. In this case, a biomarker can be used predictively to determine *ex ante* how likely a given patient is to benefit from a therapy. For example, among patients with non-small cell lung cancer, those with the ALK (anaplastic lymphoma kinase) gene mutation will benefit more from therapies like alectinib (Alecensa®) than patients without this

[7] https://www.fda.gov/Drugs/NewsEvents/ucm424545.htm

mutation. Similarly, the CFTR (cystic fibrosis transmembrane conductance regulator) modulator ivacaftor (Kalydeco®) has been approved for people with cystic fibrosis (CF) who have at least one of thirty-eight CF mutations—out of more than 1700 mutations in the gene that causes the disease. This amounts to approximately 3,500 potential patients in the United States. [8]

Many biomarkers associated with precision medicines are genomic in nature. The FDA defines a genomic biomarker as "a measurable DNA and/or RNA characteristic that is an indicator of normal biologic processes, pathogenic processes, and/or response to therapeutic or other interventions" and can be a measurement of the expression, function, or regulation of a gene (FDA, 2008). In recent years, a host of new genomics companies have sprung up, providing genetic sequencing technologies, including both software and hardware. An early 2017 report found that companies in genomics and sequencing raised more money in 2016 than any other category of digital health companies (Rock Health, 2017).

In response to the growing market and the scientific and regulatory knowledge needed to commercialize such technologies, public funding organizations and regulators have joined forces to harmonize language around biomarkers: in 2015, the joint leadership council of the FDA and NIH identified "the harmonization of terms used in translational science and medical product development…with a focus on terms related to study endpoints and biomarkers" as a priority need. One product of this effort was the publication of the "BEST ("Biomarkers, EndpointS, and other Tools) Resource" in December of 2016 (FDA and NIH, 2016). Appendix A lists the biomarker definitions established to-date by the FDA-NIH Biomarker Working group.

Yet these broad discussions about biomarkers often fail to differentiate among a diverse set of biomarker applications, each of which has different economic implications. Biomarkers can reveal useful information about disease diagnosis and prognosis, predict the treatment efficacy or toxicity of a medication, serve as markers of disease progression, and often serve as auxiliary (or so-called "surrogate") endpoints in clinical trials. Complicating matters further, some biomarkers can be used in more than one way, while others have just one known role.[9] While all of these applications of biomarkers have the potential to

[8] Since Kalydeco® (ivacaftor) was initially approved in 2012 for patients with the G551D mutation, the FDA has subsequently approved its use for patients with any 1 of 38 mutations. According to the Cystic Fibrosis Foundation, recent approvals in May 2017 and August 2017 added an estimated 900 and 600 patients in the US to the estimated 2,000 who were already eligible for treatment with ivacaftor. (https://www.cff.org/News)

[9] Biomarkers come in many types (genomic, proteomic, cellular, biochemical, structural, etc.) and can take on a range of roles (uses) in both drug development and clinical practice. These are explained below and listed in Tables 2 and 3.

shape the practice of (more) personalized medicine and improve drug development and clinical practice, only a small subset has the potential to assist in the development of precision *medicines*, those therapies targeted at specific patient populations who are more likely to benefit. It is the latter group of biomarkers – and the clinical trials driven by their use – that we specifically consider here.

A key opportunity in precision medicine is therapeutic innovation. As we improve our understanding of the genetic and cellular basis of disease, it will be possible to use genetic and protein biomarkers to classify patients into increasingly more specific subtypes where specific medicines will be more effective. In addition, biomarkers that can serve as surrogate endpoints can lead to faster clinical trials, which may influence decisions about whether to pursue treatments for specific diseases (Budish, Roin, and Williams 2015). However, the development of drugs that rely on biomarkers can also introduce challenges to the traditional clinical trial process, such as increased difficulty in trial recruitment due to smaller target patient populations. Additionally, trial design and execution can be significantly more complex when a companion diagnostic (used to identify the biomarker) needs to be approved alongside the drug (Fridlyand, et.al 2013). Regardless of the specific application, an increase in the use of biomarkers has the potential to markedly change the development and approval process for pharmaceutical innovation.

The drug development pipeline

To describe the drug development pipeline for precision medicines, we characterize all phases of development-oriented clinical trials for new drug candidates over the past twenty-two years. Clinical trials oriented towards drug development typically consist of three main phases, which commence following a manufacturer's successful completion of preclinical studies and submission of an Investigational New Drug (IND) application. Phase I is primarily designed to assess product safety and appropriate dosage. Phase I trials run for several months and typically include 20-100 healthy volunteers or individuals with the target disease. Phase II trials are much larger, enrolling up to several hundred individuals with the target disease and typically lasting between several months to two years. Phase II trials are intended to study drug efficacy and side effects. Phase III trials – usually the final stage of pre-market clinical research – are the largest, enrolling anywhere from a few hundred to a few thousand individuals with the target disease. These trials are designed to study clinical efficacy and to monitor and collect data on adverse reactions to new drugs. Sometimes also referred to as "pivotal studies," Phase III trials typically take 1-4 years to run – but

can take far longer (or shorter) depending on the normal progression of the disease studied.[10] Once Phase III results are available, manufacturers must submit a New Drug Application (NDA) or Biologics License Application (BLA) to the FDA that includes the full set of results from preclinical and clinical studies. The FDA then has up to 10 months to review the application and determine whether to grant marketing approval.[11]

The role of major pharmaceutical R&D actors

Clinical trials can be funded by private companies – both small privately-financed and large publicly-listed organizations – as well as by universities/academic medical centers, and by public actors such as the NIH. The latter has historically been more focused on early-stage research with a particular focus on basic science[12] (therefore, to the extent NIH-funded studies lead to drug development projects, one would expect NIH support to be more likely to appear in the context of earlier-stage clinical trials). This focus stems from the economic role of the NIH as not only the world's largest funder of biomedical research (with nearly $32.3 billion invested in 2016), but also a provider of public goods in the form of investments in basic research.[13]

How might we expect patterns of NIH investment to differ among PPM trials? First, markets for precision medicines may be smaller (because the biomarker segments the patient population) and thus less attractive to private actors. At the same time, however, rare diseases are known to have strong lobbies: Hegde and Sampat (2015) find that approximately 3-15% of NIH grants for rare diseases are influenced by politics, suggesting that lobbying plays a role in the allocation of public resources. It is therefore possible that there could be relatively more NIH funding of later-stage precision medicines trials in response to

[10] https://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm

[11] In recent decades, the FDA has introduced several expedited approval programs for drugs intended to treat serious conditions. "Fast Track" designation allows for frequent meetings with an FDA review team and is for drugs for which there is evidence of addressing an unmet medical need or treating an infectious disease. "Breakthrough Therapy" is for drugs that have preliminary clinical evidence indicating substantial improvement over available therapies and guarantees intensive guidance from the FDA as early as Phase I while also providing several opportunities for expedited and rolling review of results. The "Accelerated Approval" pathway is for drugs that demonstrate an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit and provides the potential for approval based on that surrogate endpoint or an intermediate clinical endpoint. "Priority Review" requires the FDA to review marketing applications within 6 months rather than 10 and is available in a number of circumstances. https://www.fda.gov/downloads/Drugs/Guidances/UCM358301.pdf

[12] https://nexus.od.nih.gov/all/2016/03/25/nihs-commitment-to-basic-science/

[13] The stated mission of the NIH is "to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability."

disease group lobbies or in order to address shortfalls in private investment in these diseases. Second, PPM trials may be more innovative and closer to the frontier of biomedical research, a fact that should increase their likelihood of being supported by a competitive research grant. On the other hand, in many cases, these trials are sponsored by for-profit companies looking to commercialize targeted therapies, which can potentially be sold at higher prices, making even small markets more financially attractive (Stern, Alexander, and Chandra, 2017). In this case, private market interest in R&D projects for PPMs may over-correct for any additional propensity for such projects to receive NIH funding. The likelihood of observing public funding in PPM trials relative to other clinical trials (conditional on phase and drug indication) is therefore an empirical question. We study the role played by each of these actors in the development of PPMs and how these roles have changed over recent decades.

3. THE ECONOMICS OF PRECISION MEDICINE

As noted above, not all biomarkers imply precision medicines. Here, we outline the economics of precision medicine to better understand how and why biomarkers are important for understanding the potential future of the pharmaceutical market. We argue that biomarkers that provide surrogate endpoints help manufacturers by speeding up clinical trials – e.g. through the use of the FDA's accelerated approval process.[14] This increase in the speed of clinical trials may provide the incentive for pharmaceutical manufacturers to target drugs for different conditions, thus potentially bringing new innovation to the market (Budish, Roin, and Williams, 2015) Conditional on approval however, such drugs may be priced lower because the evidence base for their approval was less certain.[15] At a broad level, the effect of the types of biomarkers that can be used as surrogate trial endpoints is limited to changes in the length of the drug development process (via the ability to run shorter clinical trials).[16]

In contrast, biomarkers that predict treatment benefit (by defining the subset of patients who are most appropriate for therapy) can have far reaching consequences. These include the ability to run faster trials because a therapeutic effect is easier to detect as a result of the greater putative efficacy in the indicated

[14] https://www.fda.gov/drugs/resourcesforyou/healthprofessionals/ucm313768.htm

[15] This may be particularly true, for example, in cases where precision medicines are approved based on limited data and/or surrogate endpoints. Additional evidence substantiating their benefit on actual patient outcomes is likely to be required before clinicians and health organizations adopt these medications and reimbursement levels are determined (Dzau and Ginsburg 2016).

[16] For a detailed discussion of how the use of surrogate endpoints impacts drug development incentives, see Budish, Roin, and Williams (2015).

population, but also have a tendency to change expected market sizes. Further, as we have noted elsewhere, such biomarkers could facilitate indication-based pricing, which would expand access to patients, but means that higher prices will be charged for patients who have a biomarker that indicates the drug will be most effective (Chandra and Garthwaite, 2017). The broad contours of this type of price-discrimination are illustrated through a fictional example presented in Appendix B.[17]

Biomarkers can facilitate a drug market being segmented into identifiable groups based on the expected efficacy of the product – and thus the willingness to pay for the product. In a setting where pharmaceutical manufacturers are able to charge only a single price, these subgroups allow firms to choose which patients to serve. For example, where the population receiving the least value is quite large, the manufacturer can set a low price and sell to a larger market. However, when the low-value population is quite small, the manufacturer can choose a higher price and forgo sales to lower-value patients.

Economists will note that this represents the classic monopolist's dilemma, where pharmaceutical firms must trade margin for quantity. For this reason, firms attempt to find ways to sell the same product to different customers based on their valuation – a strategy known as price discrimination. If firms develop a mechanism for charging indication-based prices, the existence of well-established, readily identifiable biomarkers will allow for price discrimination. When this is feasible, the most extreme outcome is that the manufacturer is able to capture all of the surplus as profits. Depending on the distribution of patients, this could (but need not) expand access to lower-value indications. However, an indication-based pricing strategy weakly increases the profits of firms developing precision medicines. As a result, the expanded use of biomarkers has the potential to provide additional incentives to develop products that otherwise would not be commercialized. Pricing aside, biomarkers that predict treatment efficacy reduce market size, which in turn, can reduce some of the incentives for innovation. Some biomarkers will allow manufacturers to quality for "orphan drug" designation through the Orphan Drug Act of 1983 (ODA) by carving out an indication that affects fewer than 200,000 patients. If a medicine receives FDA approval for an orphan condition it receives tax credits equaling 50% of clinical trials expenses and seven years of marketing exclusivity (two years longer than non-orphan drugs). These incentives have been shown to be powerful:

[17] This figure depicts the monetary value of a hypothetical product for three different indications (for example, patient populations defined by the presence of biomarkers), the size of the patient populations affected by each indication, and the prices charged for the product under different pricing regimes.

more than 516 medicines for over 450 different rare diseases have been approved through the ODA[18] and in 2015 alone, 47% of novel drugs approved were orphan drugs.[19] When such an approval happens, it will also raise prices because of the (extended) protections from generic competition offered by the Orphan Drug Act, and the fact that smaller markets will attract less follow-on competition. Further, in small markets, brand-brand competition will be far less robust than in large markets. Thus, even after exclusivity periods end, there may not be a substantial enough market to stimulate price competition through generic entry.

Finally, the complexity of developing products in this space combined with the use of new and emerging technologies may result in greater specialization for the drug development process. This could involve a greater share of products beginning their lifecycle at small research-focused firms than would be true in more traditional segments of the pharmaceutical industry.

4. DATA

We use data from the Cortellis Competitive Intelligence Clinical Trials Database (Cortellis), which is compiled by Clarivate (and formerly by Thompson Reuters). The database includes over 270,000 global and US-based clinical trials. Cortellis includes full coverage of 24 clinical trial registries from around the world, including clinicaltrials.gov, which is maintained by the National Institutes of Health (NIH), and European Clinical Trials Database (EudraCT), which is maintained by the European Medicines Agency (EMA). Biomedical researchers are strongly encouraged to register trials for publication in medical journals and, as of 2005, trials must be registered to an approved public clinical trial registry prior to patient enrollment in order to be considered for publication in any International Committee of Medical Journal Editors (ICMJE) member journals (De Angelis, Catherine, et al., 2004).

Because both publication and registration are integral parts of the new drug development process, the set of registered trials included in Cortellis should capture all relevant development efforts – in particular, in the years since 2005.[20] Cortellis has full coverage of all ICMJE approved trial registries (Clinical Trial

[18] https://www.accessdata.fda.gov/scripts/opdlisting/oopd/index.cfm
[19] http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm474696.htm
[20] We believe that coverage of registered trials is comprehensive and we further expect a high share of trials to be registered in the post-2004 period (De Angelis, Catherine, et al., 2004). However, we note that certain types of trials – e.g. smaller

Registration, 2016) and Cortellis data have been used in several published studies in peer reviewed biomedical journals such as Lancet Infectious Disease (Phyo, Aung Pyae, et al., 2016) and Nature Reviews Drug Discovery (Bespalov, Anton, et al., 2016). Appendix C includes a detailed timeline of important dates related to the registration of clinical trials and the establishment of the U.S. registry *clinicaltrials.gov*.

Data composition and summary statistics

We queried the Cortellis database for all clinical trials with a launch date between January 1, 1995 and December 31, 2016 for a total of 22 calendar years of clinical trial starts. We identify the full set of phase I, II, and III[21] clinical trials, along with detailed clinical trial information associated with each trial. A few facts are notable: first, the total number of registered trials worldwide has grown over time for each phase of clinical research (Figure 1) and in particular for phase II trials.[22] In 2016, roughly six thousand phase II trials were launched globally, nearly double the number of registered trials launched a decade earlier, in 2006.

For each trial, the Cortellis database also provides information on the trial's relevant drug indication(s), any biomarkers used in the trial, and the trial's sponsor(s). In addition, we are able to classify trials according to broad set of descriptive categories such as whether a product is a small-molecule drug or a biologic, and the presence (or absence) of one or more biomarker(s) used in the trial. For each biomarker, we are separately able to consider data on its type and use (role). A complete list of the descriptive variables we consider and their frequencies in the clinical trials data set are provided in Table 1.

trials without regulatory oversight may still be missing in our data. Kao (2017), describes these types of trials and how they may be designed to signal "off-label usability" to physicians. While an understanding of these types of unregistered trials is important for understanding pharmaceutical firm strategy, we do not believe they are likely to be the types of trials that we attempt to identify in this study, which are those specifically intended to commercialize targeted therapies.

[21] For the simplest classification of trails into phases, we assign combined trials (e.g. combined Phase II & Phase III) to the lower of the two phases involved. For example, a combined Phase II/ Phase III trial would be classified as having started Phase II in the year that the trial launched. In robustness tests, we create separate sub-categories for combined Phase I/II and II/III trials and include controls for these combined trials in regression analyses. Subsequent regression results are not sensitive to this distinction, so we use the simplified 3-phase classification in tables and figures for simplicity.

[22] The recent spike in the number of global clinical trials (and Phase II trials in particular) is driven by growth in non-US trails (see Appendix tables for a version of Figure 1 that presents only U.S. trials).

To aggregate the detailed indications reported in the Cortellis database to more usable categories, we used a dataset[23] of indications matched to ICD-9 codes to link each trial in our dataset to a 3-digit ICD-9 code. The matched indication-ICD-9 dataset was independently checked for accuracy by two research assistants using an online ICD-9 medical coding reference manual,[24] and any discrepancies between their matches were resolved by a third research assistant. Each trial was ultimately assigned to one ICD-9 code, corresponding to a total of 65 ICD-9 sub-chapters (Appendix D).

We also capture key information about the clinical trial's sponsor(s). Trial sponsors are identified by name and type, including academic investigators, government, non-government, company, and other sponsors. Importantly, the database also lists associated clinical trial registry identifiers such as unique trial registration numbers from clinicaltrials.gov as well as NIH grant numbers that supported the research. We use these to identify whether a trial benefited from NIH funding and, if so, the type(s) of NIH grant(s) listed as a funding source. NIH funded trials can be segmented by type of NIH funding using the activity code embedded in the NIH project number(s) listed. Appendix E describes how NIH project types are identified.

The categorical variable "biomarker type" indicates the biological feature that a given biomarker identifies. Biomarker types include genomic biomarkers, proteomic biomarkers, biochemical biomarkers, cellular biomarkers, physiological biomarkers, structural biomarkers, and anthropomorphic biomarkers. Definitions of biomarker types and their frequencies of use in clinical trials both a) overall and b) over time are reported in Table 2. Importantly, these types are not mutually exclusive, since a given biomarker – e.g. a receptor such as EGFR (epidermal growth factor receptor) – can be both a genomic and proteomic biomarker. This is because genomic characteristics will lead to differential expression of EGFR – making it a biomarker of particular genomic features –but EGFR is *itself* a protein and therefore a proteomic biomarker as well. For this reason, there can be correlation in the frequencies of biomarkers types across trials.

Biomarker data and defining pipeline precision medicines

[23] We are grateful to Manuel Hermosilla, Craig Garthwaite, and David Dranove, who generously shared their version of a 3-digit ICD-9 crosswalk dataset with us. This dataset was assembled through the use of two independent medical coders separately constructing a crosswalk. Discrepancies were adjudicated by a third expert and additional outside research.
[24] ICD9Data.com

Cortellis includes fairly broad categories of biomarker uses as they may relate to clinical trials. These include disease markers, toxic effect markers, and therapeutic effect markers. Disease-related biomarkers indicate if a disease already exists (diagnostic biomarker), or how such a disease may develop in an individual case regardless of the type of treatment (prognostic biomarker). Therapeutic effect-related biomarkers provide an indication of the progress of a product on the patient during treatment. Toxic effect-related biomarkers indicate a treatment-related adverse reaction. Other biomarker roles are "not determined" because they do not have any of the roles described above in a particular trial. In practice, we are interested in a *subset* of the trials that use disease-related biomarkers – namely, those in which we observe the unambiguous features of products that would likely come to market as targeted therapeutics upon successful progression through the R&D process. This is because this subset of biomarkers facilitates *ad hoc* patient selection for therapy.

Our working definition of potential precision medicines (PPMs) is that they encompass the set of pipeline products that are being developed using the types of diseases-related biomarkers that are relevant for identifying subpopulations that are likely to be more (or less) responsive to medications. We therefore employ a second, biomarker-specific database from Clarivate in order to link biomarkers to their *detailed* roles in clinical trials. The detailed biomarkers data (DBD) from Clarivate include additional detail (in the form of "detailed biomarker roles") on all known clinical biomarkers and their paired uses and indications in clinical research. For example, human epidermal growth factor receptor 2 (HER2) is a (genomic) biomarker that can be used for (a) selection for therapy and (b) predicting treatment efficacy – both of which are detailed biomarker roles – among patients with breast cancer (the indication). Based on using a trial's "breast cancer" indication and knowing that the HER2 biomarker was used in that clinical trial, one can probabilistically assign both a biomarker type and a detailed biomarker role (or, in some cases, more than one) to that trial. Generally, in order to assign biomarker types and biomarker roles to our full set of clinical trials, we match the named biomarker(s) associated with each trial (when there are any) and the indication of that trial to the DBD.

Definitions of detailed biomarker roles and the frequencies of their use in clinical trials are reported in Table 3. A biomarker may have multiple associated uses, making it important to correctly link a biomarker associated with a given clinical trial and indication to its use *in that setting*. Therefore, the process of matching a biomarker-indication pair from the Cortellis clinical trials data with a biomarker-indication pair from

the DBD is a crucial step in correctly assigning biomarker roles to individual clinical trials. We define PPMs in two ways using these detailed biomarker roles. These classifications are consistent with the NIH-FDA definitions of biomarkers and how they are employed and were separately discussed with an oncologist, who is the principal investigator on a biomarker-driven clinical trial.

In the first, "generous" definition of PPMs, we identify trials using biomarkers whose roles include diagnosis, differential diagnosis, predicting drug resistance, predicting treatment efficacy, predicting treatment toxicity, screening, and selection for therapy. The rationale for the generous definition is that all of these biomarkers can be used in the development of targeted therapeutics and are likely to be associated with the development of precision medicines.

In the second, "restrictive" definition of PPMs, we identify the subset of trials from the "generously" defined group that specifically employ biomarkers for prediction (predicting drug resistance, predicting treatment efficacy, and predicting treatment toxicity), with the vast majority of these trials identified as PPM trials due to the use of biomarkers that can help predict treatment efficacy (Table 3). We consider each in turn and further consider the interaction of these roles with specific biomarker *types* (genomic and proteomic) that are more likely to be used in trials for precision medicines.

5. CHARACTERIZING THE PPM DEVELOPMENT PIPELINE

We characterize the number and type of drugs using biomarkers in their clinical trials as well as those that can be considered PPMs by therapeutic area and over time. Since cancers represent a set of diseases in which precision therapies are already successfully used, and since cancer applications of precision medicine are expected to grow in coming years, we separately characterize the cancer applications of pipeline precision medicines in detail.

Biomarkers and PPMs in clinical trials

We begin at perhaps the broadest point, by first identifying all trials that use one or more biomarker(s) of any kind (Figure 2). Notably, both the share and total number of clinical trials employing biomarkers has increased markedly over recent decades. We next focus only on the subset of biomarker uses that are associated with PPMs, by both the generous and restrictive definitions (Table 3). Both the number and

percentage of PPM trials increased over our period of observation, as seen in Figure 3. We further note that the two definitions of PPMs track each other closely over time – both in Figure 3 as well as in the subsequent sub-sample analyses described below. Table 4 presents the count (column 1) and percentage (column 2) of PPMs in clinical trials in each year of our data. Columns 3-8 present the same results by clinical trial phase. Even by the most restrictive definition of PPM trials, by 2016, approximately 8.7% of trials were for PPMs, more than double the percentage a decade earlier in 2006 (4.2%).

PPMs were associated with different types of biomarkers and the relative and absolute frequencies of these types have evolved over time. Biomarker types are not mutually exclusive; for example, there is near-complete overlap between proteomic and genomic biomarkers, since the vast majority of genomic mutations (e.g. in cancer) manifest themselves through differences in protein expression. Figure 4 shows how these types were represented in each phase (by both generous and conservative definitions of PPMs), over our years of observation. Genomic/proteomic biomarkers were the most commonly used in recent years, featured in the vast majority of PPM trials, a statistic that is consistent with PPMs being driven primarily by understanding gene and protein expression and how these factors predict the likely success of medications.

Pipeline precision cancer therapies

Figure 5 and Table 5 present data on the frequency of PPMs in cancer trials only. Several features of these trials are notable – especially in comparison. First, PPM trials are more than an order of magnitude more common in cancer indications: in 2015 and 2016, roughly 25% (or more) of all cancer drug trials were PPM trials, but only 1-2% of trials for non-cancer indications were PPM trials. In regression analysis, we also see that a cancer indication is a strong statistical predictor of a PPM trial and the growth of PPMs among cancer drugs explains the lion's share of growth in PPM trials over the past two decades. These results are completely consistent with the clinical belief that the majority of applications of precision medicines in coming years will be in the context of targeted therapies for cancer.

Institutional factors

Next, we consider the PPM development pipeline in light of a number of specific institutional factors. We consider US-based vs. non-US-based trials. The United States is, by far, the world's largest pharmaceutical consumer (International Trade Administration, 2016) and it would therefore be reasonable to expect trials for PPMs to be driven to by both local demand (Costinot et al., 2016) as well as local regulations (FDA, 2004). Figure 6 shows the number and share of U.S. PPM trials. The total number of PPM trials conducted within the United States is comparable to the total number conducted abroad, but the share of PPM trials among U.S. trials is roughly double that of international trials. This finding is consistent with the fact that U.S. drug prices are typically higher than those of other countries (Kanavos et al., 2013), making it more appealing for pharmaceutical manufacturers to bring drugs to market in the United States as soon as possible. These facts are also reflected in our regression analysis which shows that U.S. trials are 0.5-0.6% more likely to involve PPMs at any point in time than their non-U.S. counterparts in the same year, all else equal.

We next consider PPM trials with vs. without NIH funding. Since NIH grants are concentrated in U.S. research institutions, we focus our analysis of NIH funding on U.S. trials only. The first panel of Figure 7 shows the share of clinical trials by phase that received NIH funding in each year. Although NIH funding has grown over time, the total number of registered clinical trials has grown more rapidly, leading to a declining *share* of trials acknowledging NIH funding. Among PPM trials (both definitions), the share of trials with NIH funding has been relatively constant, albeit somewhat noisily measured. On average, roughly 5-6% of Phase I and Phase II trials (but a lower share of Phase III trials) have received NIH support in recent years (with higher averages, but also higher variances in earlier years of observation; Figure 9). Table 6 presents the absolute shares of all trials – not restricting to PPM trials – receiving NIH funding over our sample. Overall, NIH support is *less* common among PPM trials (of both definitions) relative to overall rates of NIH support of clinical trials in the United States.

Finally, we consider the types of firms (publicly listed vs. small/private) engaging in the development of PPMs and conclude with regression analysis (Tables 7-8). We are circumspect in interpreting our regression results: the coefficients calculated are not *causally* estimated; rather they represent differences between categories in our sample, controlling for other factors. However, the coefficients are useful in that they allow for interpretation of multivariate associations. For example, the linear probability models presented in Table 7 (Appendix Table III presents marginal effects from logit specifications) indicate that the

share of PPM trials has been increasing over time by roughly 0.4 percentage points per year. Other relationships seen in earlier tables and figures are also apparent. Most prominent among these is the importance of cancer trials: trials for cancer indications are 14-15 percentage points more likely to be PPM trials than those for non-cancer indications. Trials with U.S. sites are also more likely than non-U.S. trials to be PPM trials, but only by about one percentage point – in other words, U.S. trials seem to be about 2.5 years "ahead" of non-U.S. trials in their inclusion of PPMs.

6. CONCLUSION

By taking a detailed view of the global clinical trial pipeline, we consider several trends in clinical trials over recent decades. Beyond growth in the number of registered clinical trials, we document a number of patterns that that have implications for cost-growth in health care and pharmaceutical pricing. First, we document that the use of biomarkers in clinical trials has grown significantly, with an important subset of those representing the types of biomarkers that have the potential to be used in the development of targeted therapies. Such therapies are likely to be more effective, but will also likely come with higher prices. Although the raw numbers of trials using biomarkers in the development of precision medicines is still dwarfed by the total number of clinical trials, the growth in such trials has been large in percentage terms, approximately doubling every decade over the past 22 years.

Our results should be interpreted with a number of caveats. Firstly, the findings presented here are only as representative as the global registries on which our primary clinical trial dataset is based. While we have noted above that there are good reasons to believe that these registries are highly representative of the set of pipeline drugs pursuing regulatory approvals in the dozen most recent years of our data, some trials may not have been reported in earlier years. Unfortunately, we do not have a data-driven way of estimating the type and direction of selection into trial registries that may have occurred.

Secondly, we note that our characterization of trials as either PPM or non-PPM trials is, by nature, probabilistic, based on observable features of these trials and the drugs in them. While the categories we use are unambiguously more conservative than simply considering any use of biomarkers in clinical trials, they may still capture some trials and pipeline products that do not, in fact, represent true PPMs.

Finally, and perhaps most importantly, we have characterized the drug development *pipeline*, which is not necessarily synonymous with characterizing the *actual set of products* that are commercialized. If failure

rates in clinical research are endogenously determined with other characteristics related to commercialization strategies, it is easy to imagine how the set of products that are commercialized might, on average, look different than the late-stage clinical trial pipeline would suggest.[25] Indeed, within the pipeline itself, there is evidence that firm characteristics are correlated with the probability that a product is advanced into the next stage of clinical research (Guedj and Scharfstein, 2004).

Yet we believe that we have also made progress in characterizing recent trends and developments in clinical research related to precision medicines. By taking a big-picture view of clinical trials globally, we can observe how PPMs have grown in number and share of trials over recent decades. We can also bring empirical data to bear on predictions from medicine and economics, which would suggest that certain types of drugs (e.g. for cancers) and certain markets (e.g. in the United Sates) are likely to have a greater share of PPMs. Within PPMs, it seems like drug development is being driven by genomic/proteomic biomarkers, suggesting the growing importance of sequencing technologies for both R&D and patient care. Further, recent trajectories have implications for health care spending: to the extent that PPMs grow in market share, they will drive up costs for drugs that target specific groups of patients and also open up opportunities for indication-based pricing.

[25] On average, success rate for a drug entering clinical trials is approximately 10%. This rate is even lower for oncology therapeutics at roughly 5%. (https://www.bio.org/sites/default/files/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf)

Figures

Figure 1: Clinical trials over time

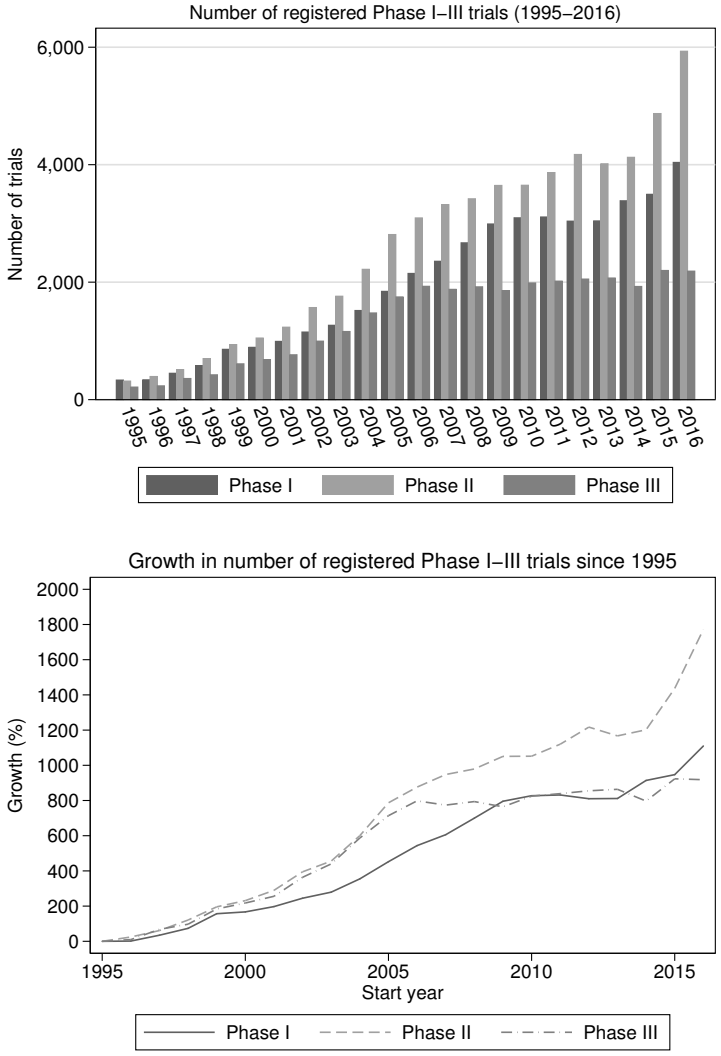


Figure 2: Clinical trials using biomarkers

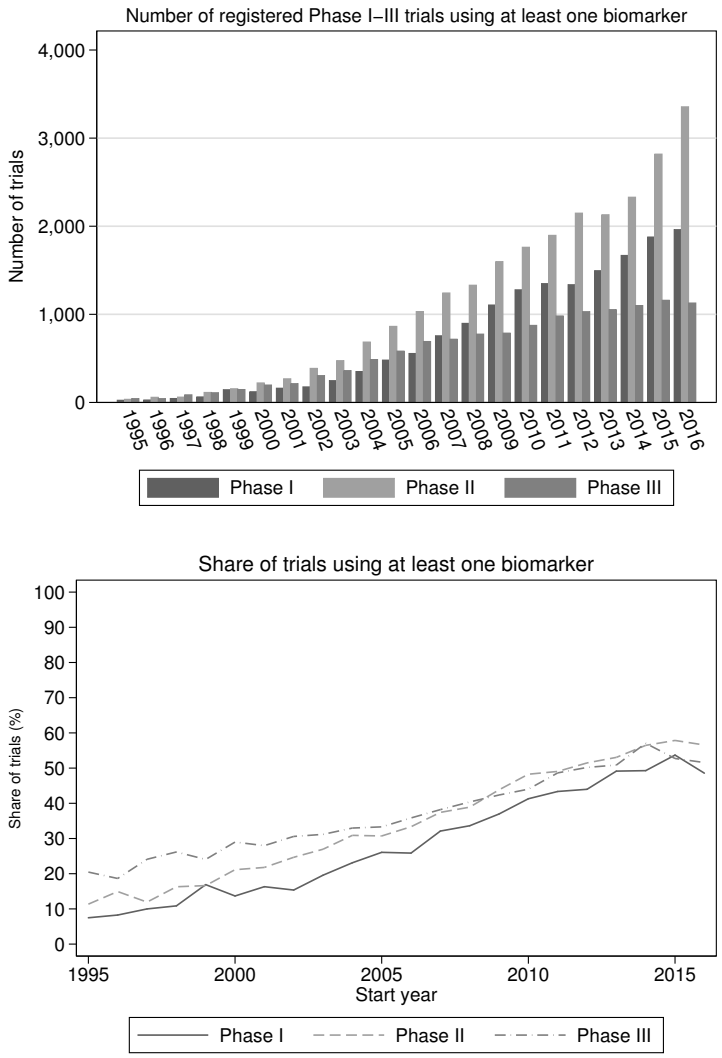


Figure 3: Clinical trials for PPMs

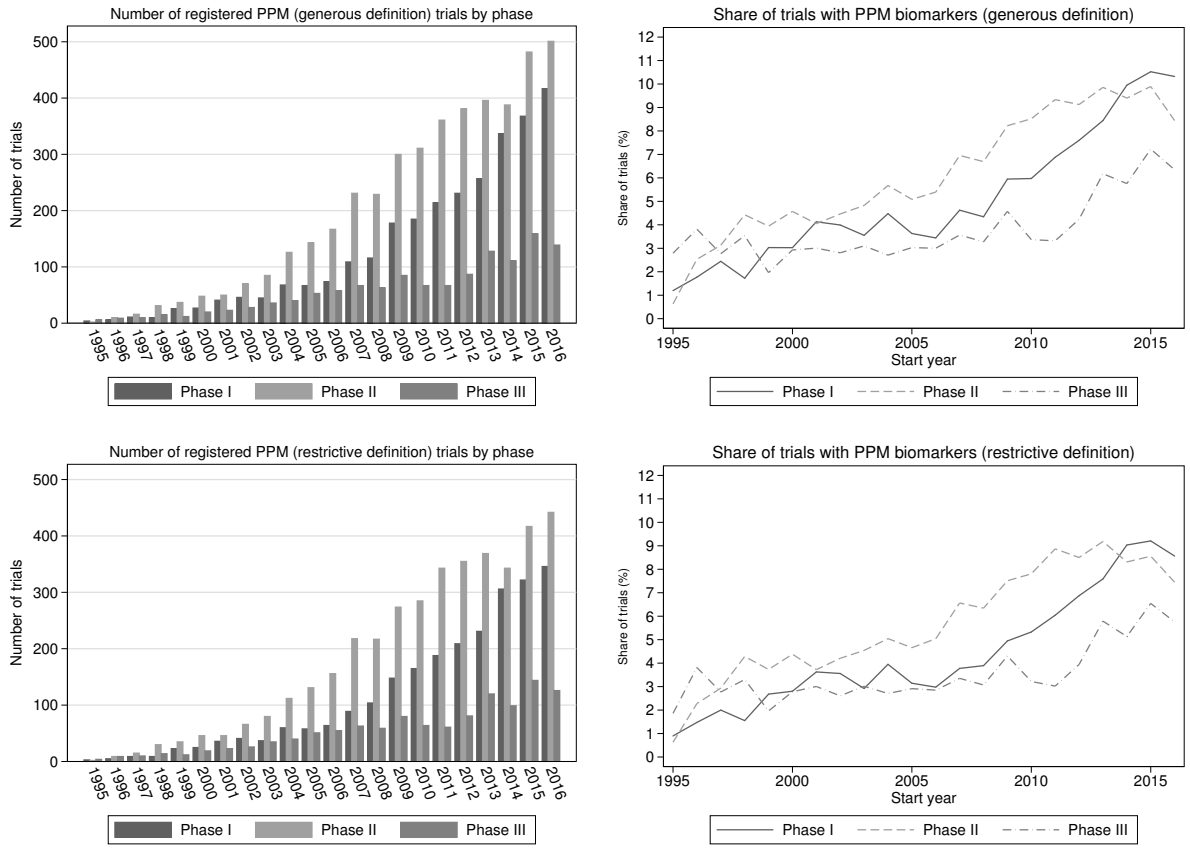


Figure 4: Types of biomarkers used in PPM trials

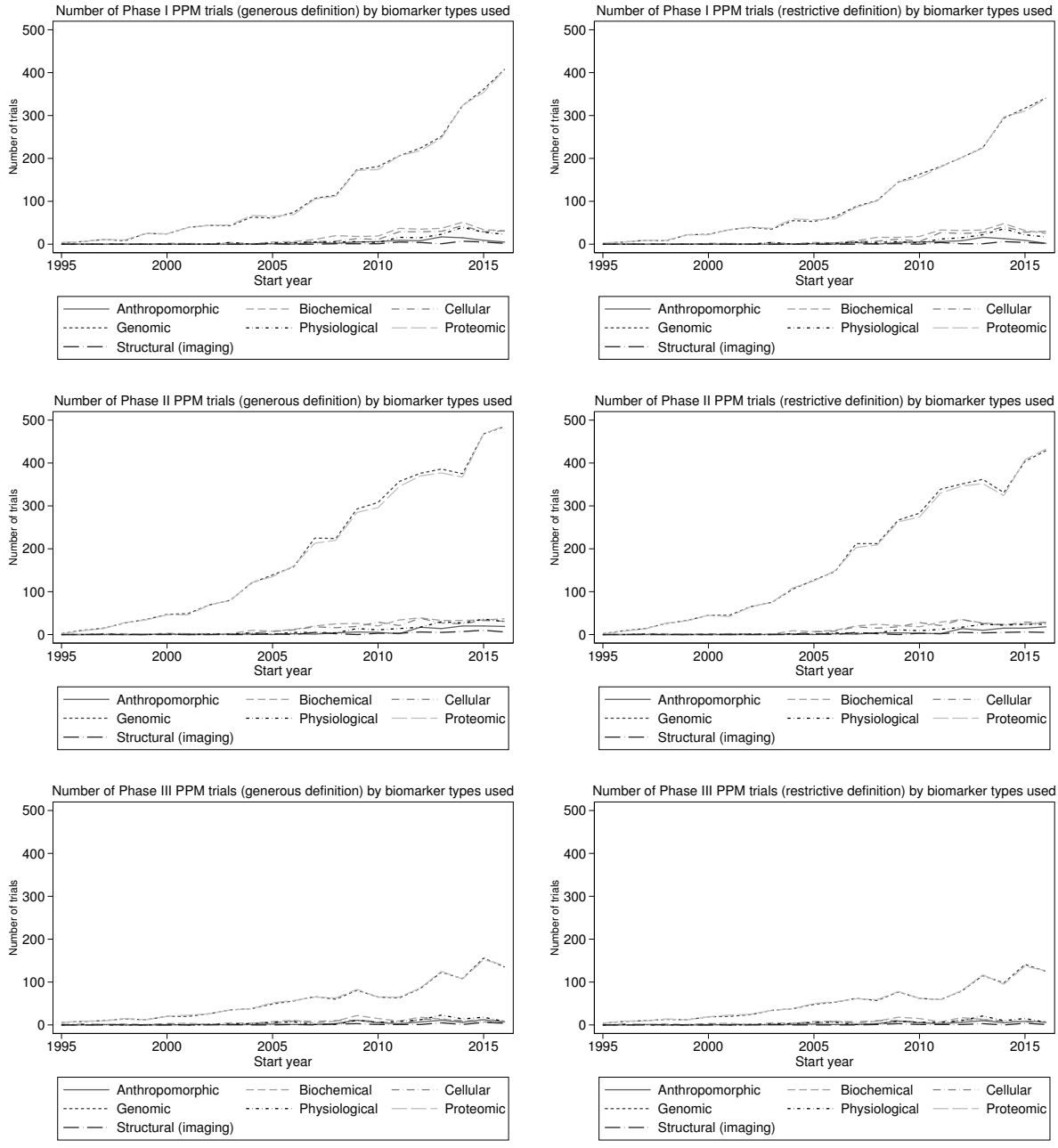


Figure 5: Clinical trials for PPMs, cancer indications only

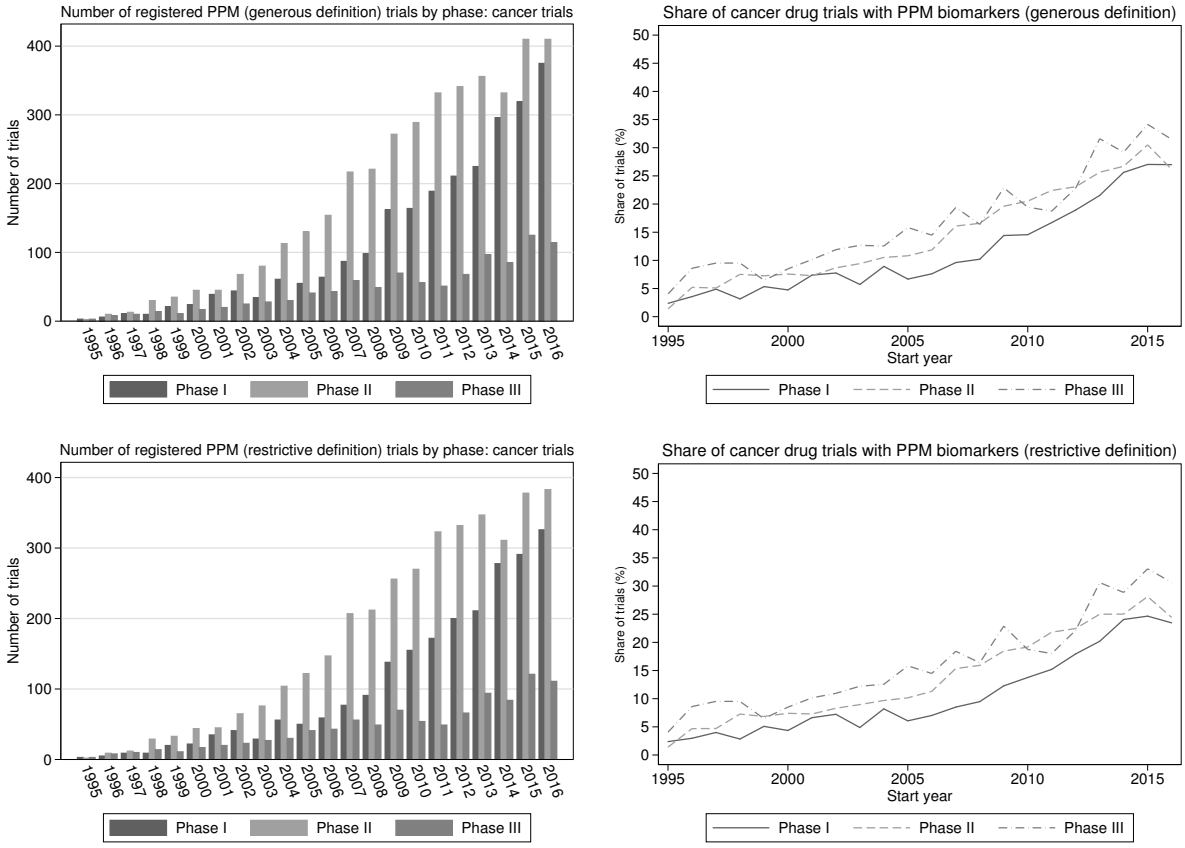


Figure 6: Clinical trials for PPMs, U.S. trials only

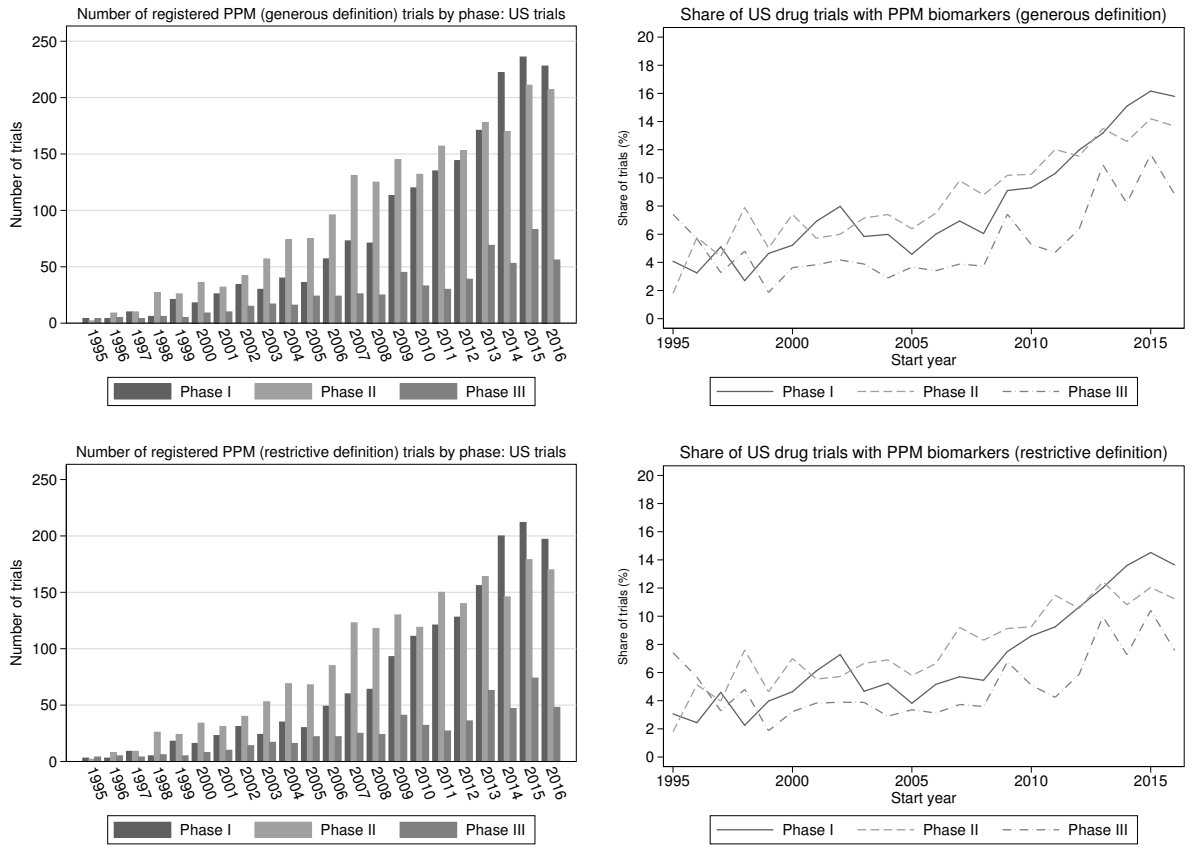
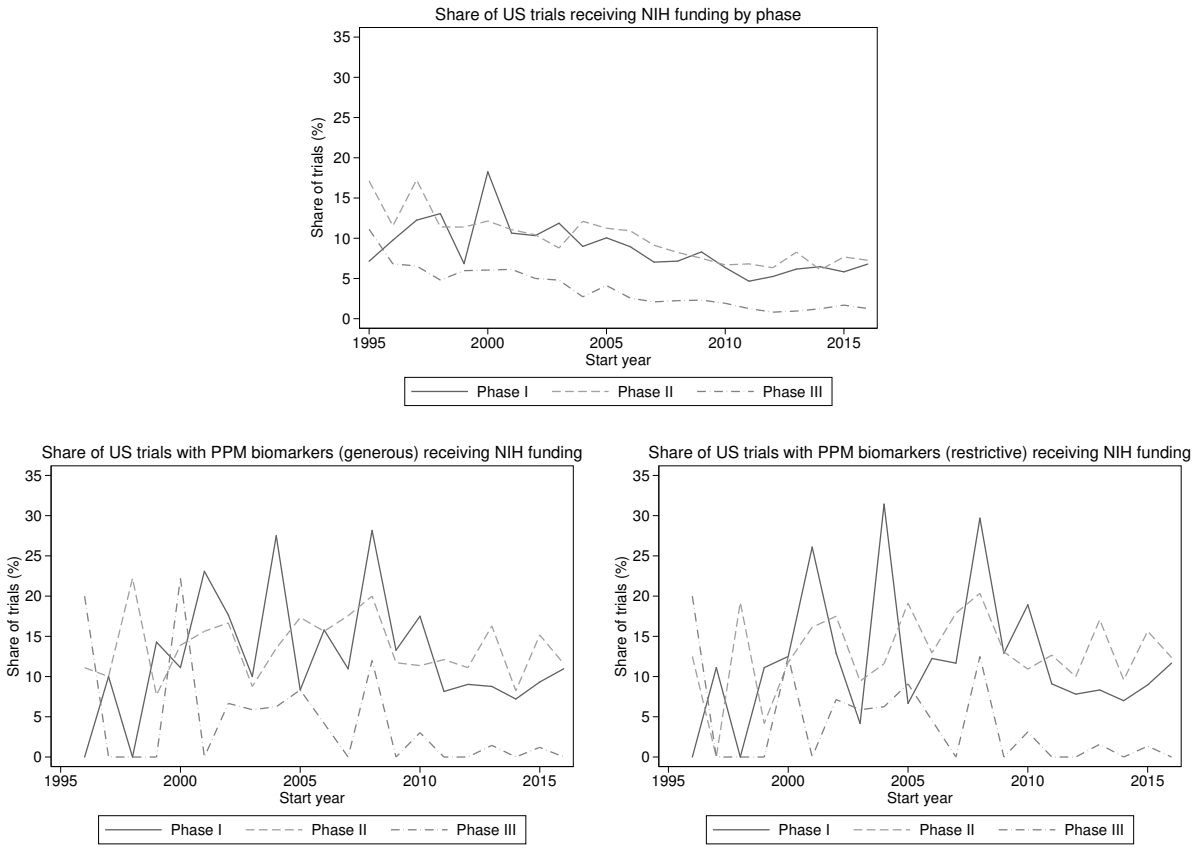


Figure 7: Trials and PPM trials with NIH funding (U.S. trials only)



Tables

Table 1: Summary statistics for selected variables

	All trials		US trials	
	Mean	Observations	Mean	Observations
Uses biomarker	0.4092	131,971	0.4619	49,540
Generous PPM	0.0643	131,971	0.0907	49,540
Restrictive PPM	0.0581	131,971	0.0813	49,540
Biomarker role: disease	0.0842	131,971	0.1145	49,540
Biomarker role: toxic effect	0.0496	131,971	0.0699	49,540
Biomarker role: therapeutic effect	0.3371	131,971	0.3758	49,540
Biomarker role: not determined	0.0023	131,971	0.0024	49,540
Biomarker type: anthropomorphic	0.0350	131,971	0.0400	49,540
Biomarker type: biochemical	0.1248	131,971	0.1300	49,540
Biomarker type: cellular	0.0308	131,971	0.0424	49,540
Biomarker type: genomic	0.2321	131,971	0.2845	49,540
Biomarker type: physiological	0.0849	131,971	0.0865	49,540
Biomarker type: proteomic	0.2426	131,971	0.2942	49,540
Biomarker type: structural (imaging)	0.0177	131,971	0.0200	49,540
Biomarker role (detailed): diagnosis	0.2948	117,180	0.3448	43,777
Biomarker role (detailed): differential diagnosis	0.1829	117,180	0.2041	43,777
Biomarker role (detailed): predicting drug resistance	0.0624	117,180	0.0778	43,777
Biomarker role (detailed): predicting treatment efficacy	0.2568	117,180	0.3060	43,777
Biomarker role (detailed): predicting treatment toxicity	0.0474	117,180	0.0493	43,777
Biomarker role (detailed): screening	0.0523	117,180	0.0547	43,777
Biomarker role (detailed): selection for therapy	0.0938	117,180	0.1111	43,777
Phase 1 Clinical	0.3305	131,971	0.3653	49,540
Phase 2 Clinical	0.4367	131,971	0.4263	49,540
Phase 3 Clinical	0.2328	131,971	0.2083	49,540
Drug indication for neoplasm	0.3340	131,971	0.3990	49,540
Received NIH funding	0.0282	131,971	0.0703	49,540
Trial site in US	0.4368	113,410	1.0000	49,540
N	131,971		49,540	

Table 2: Number of trials employing biomarkers by type

	Any biomarker	Anthropomorphic	Biochemical	Cellular	Genomic	Physiological	Proteomic	Structural
Overall	53,998	4,620	16,472	4,070	30,634	11,205	32,011	2,340
1995	105	4	29	1	59	22	60	4
1996	131	5	34	6	77	16	84	4
1997	193	10	62	8	119	24	125	2
1998	288	12	74	6	165	58	182	5
1999	448	16	119	22	292	68	307	8
2000	542	33	149	28	349	83	360	9
2001	645	36	190	38	406	94	426	9
2002	869	53	263	36	558	135	579	21
2003	1,085	80	358	51	698	156	732	28
2004	1,524	126	469	68	950	216	997	34
2005	1,928	135	580	118	1,157	314	1,218	58
2006	2,280	178	737	138	1,379	377	1,462	73
2007	2,718	220	831	207	1,687	437	1,751	98
2008	3,005	252	970	245	1,813	548	1,900	101
2009	3,492	288	1,137	251	2,157	627	2,248	114
2010	3,916	334	1,239	304	2,333	740	2,418	134
2011	4,228	366	1,353	357	2,525	828	2,638	164
2012	4,517	408	1,463	406	2,566	994	2,661	206
2013	4,681	439	1,446	382	2,544	1,104	2,666	241
2014	5,099	518	1,576	434	2,647	1,310	2,762	270
2015	5,857	546	1,610	438	2,944	1,499	3,086	374
2016	6,447	561	1,783	526	3,209	1,555	3,349	383

Biomarker types:
Anthropomorphic biomarkers are markers of the body shape/form
Biochemical biomarkers are substrates or products of chemical reactions in the body
Cellular biomarkers are whole cells
Genomic biomarkers are variants in the DNA sequence or in the transcription level;
Physiological biomarkers are body processes
Proteomic biomarkers are variants in protein sequence, protein levels in a given tissue, protein interactions and enzyme activities
Structural biomarkers are anatomical structures

Table 3: Number of trials employing biomarkers by detailed role

	Any biomarker	Diagnosis	Differential Diagnosis	Predicting drug resistance	Predicting treatment efficacy	Predicting treatment toxicity	Screening	Selection for therapy
Overall	39,207	34,545	21,429	7,312	30,091	5,556	6,133	10,988
1995	105	68	45	7	62	8	8	13
1996	131	88	49	22	81	14	14	19
1997	193	130	83	38	122	31	39	38
1998	288	210	137	66	199	53	42	68
1999	448	341	201	76	310	85	57	97
2000	542	369	233	88	343	78	59	118
2001	645	458	275	121	421	85	81	138
2002	869	624	395	151	578	122	109	203
2003	1,085	764	487	174	691	157	132	263
2004	1,524	1,051	675	240	954	224	190	332
2005	1,928	1,306	799	286	1,189	263	239	408
2006	2,280	1,575	1,004	370	1,396	308	291	510
2007	2,718	1,882	1,215	444	1,693	369	332	617
2008	3,005	2,046	1,360	496	1,832	430	362	661
2009	3,492	2,352	1,578	649	2,145	504	482	842
2010	3,916	2,539	1,540	581	2,210	343	444	768
2011	4,228	2,738	1,698	582	2,379	376	502	890
2012	4,517	2,909	1,780	574	2,494	376	462	906
2013	4,681	2,932	1,778	609	2,530	396	500	964
2014	5,099	3,071	1,809	548	2,552	409	519	934
2015	5,857	3,355	2,005	574	2,816	427	589	1,070
2016	6,447	3,737	2,283	616	3,094	498	680	1,129

Biomarker roles (uses) that are related to the development of PPMs, generously defined, are included above
The restrictive definition of PPMs limits the definition to those related only to prediction: predicting drug resistance, treatment efficacy, and treatment toxicity and is driven by "predicting treatment efficacy."
Biomarker roles (uses) that are unrelated to developing PPMs, but included in the data are: disease profiling, monitoring disease progression, monitoring treatment efficacy, monitoring treatment toxicity, prognosis, prognosis - risk stratification, risk factor, staging, and toxicity profiling

Table 4: Potential precision medicine (PPM) trials (1995-2016):

	Generous definition							
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	12	1.39	4	1.2	2	.631	6	2.79
1996	25	2.58	6	1.77	10	2.53	9	3.81
1997	37	2.8	11	2.44	16	3.13	10	2.77
1998	56	3.29	10	1.72	31	4.43	15	3.54
1999	75	3.12	26	3.03	37	3.94	12	1.96
2000	95	3.62	27	3.03	48	4.57	20	2.93
2001	114	3.81	41	4.13	50	4.05	23	3.01
2002	144	3.87	46	3.99	70	4.46	28	2.81
2003	166	3.96	45	3.55	85	4.82	36	3.1
2004	234	4.49	68	4.48	126	5.68	40	2.71
2005	263	4.1	67	3.63	143	5.09	53	3.03
2006	299	4.17	74	3.44	167	5.4	58	3
2007	407	5.39	109	4.62	231	6.96	67	3.57
2008	408	5.09	116	4.34	229	6.69	63	3.28
2009	563	6.63	178	5.95	300	8.22	85	4.57
2010	563	6.44	185	5.97	311	8.52	67	3.37
2011	642	7.14	214	6.88	361	9.34	67	3.32
2012	699	7.54	231	7.6	381	9.13	87	4.24
2013	781	8.55	257	8.44	396	9.86	128	6.18
2014	836	8.85	337	9.95	388	9.4	111	5.76
2015	1,009	9.55	368	10.5	482	9.89	159	7.23
2016	1,057	8.69	417	10.3	501	8.44	139	6.35
	Restrictive definition							
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	9	1.04	3	.898	2	.631	4	1.86
1996	23	2.37	5	1.47	9	2.28	9	3.81
1997	34	2.57	9	2	15	2.94	10	2.77
1998	53	3.11	9	1.55	30	4.29	14	3.3
1999	70	2.91	23	2.68	35	3.73	12	1.96
2000	90	3.43	25	2.8	46	4.38	19	2.78
2001	105	3.51	36	3.63	46	3.72	23	3.01
2002	133	3.58	41	3.56	66	4.21	26	2.61
2003	152	3.63	37	2.92	80	4.54	35	3.01
2004	212	4.06	60	3.95	112	5.05	40	2.71
2005	240	3.75	58	3.14	131	4.66	51	2.91
2006	275	3.83	64	2.98	156	5.04	55	2.85
2007	370	4.9	89	3.78	218	6.56	63	3.35
2008	380	4.74	104	3.89	217	6.34	59	3.07
2009	502	5.91	148	4.95	274	7.51	80	4.31
2010	514	5.88	165	5.33	285	7.81	64	3.22
2011	592	6.58	188	6.04	343	8.87	61	3.02
2012	645	6.96	209	6.88	355	8.5	81	3.94
2013	720	7.88	231	7.59	369	9.19	120	5.79
2014	748	7.92	306	9.03	343	8.31	99	5.13
2015	883	8.35	322	9.21	417	8.56	144	6.55
2016	914	7.52	346	8.56	442	7.45	126	5.76

Table 5: Potential precision medicine (PPM) trials: cancer only (1995-2016):

Generous definition								
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	8	2.33	3	2.38	2	1.4	3	4.05
1996	24	5.3	6	3.57	10	5.21	8	8.6
1997	34	5.81	11	4.89	13	5.1	10	9.52
1998	54	6.26	10	3.15	30	7.52	14	9.52
1999	67	6.41	21	5.34	35	7.26	11	6.47
2000	86	6.62	24	4.75	45	7.58	17	8.5
2001	104	7.74	39	7.39	45	7.28	20	10.2
2002	137	8.78	44	7.76	68	8.68	25	11.9
2003	142	8.53	34	5.72	80	9.42	28	12.7
2004	204	10.2	61	8.93	113	10.5	30	12.6
2005	226	9.89	55	6.67	130	10.8	41	15.8
2006	261	10.7	64	7.62	154	11.8	43	14.5
2007	363	14.2	87	9.61	217	16.1	59	19.4
2008	368	14.2	98	10.2	221	16.6	49	16.4
2009	504	17.9	162	14.4	272	19.6	70	22.9
2010	509	18	164	14.6	289	20.5	56	19.4
2011	572	19.8	189	16.7	332	22.4	51	18.8
2012	620	21.5	211	18.9	341	23.1	68	22.8
2013	678	24.7	225	21.5	356	25.6	97	31.6
2014	713	26.5	296	25.6	332	26.7	85	29.2
2015	854	29.6	319	27	410	30.5	125	34.2
2016	899	27.1	375	27	410	26.2	114	31.5
Restrictive definition								
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	8	2.33	3	2.38	2	1.4	3	4.05
1996	22	4.86	5	2.98	9	4.69	8	8.6
1997	31	5.3	9	4	12	4.71	10	9.52
1998	52	6.03	9	2.84	29	7.27	14	9.52
1999	64	6.12	20	5.09	33	6.85	11	6.47
2000	83	6.39	22	4.36	44	7.41	17	8.5
2001	100	7.45	35	6.63	45	7.28	20	10.2
2002	129	8.27	41	7.23	65	8.3	23	11
2003	132	7.93	29	4.88	76	8.95	27	12.2
2004	190	9.52	56	8.2	104	9.68	30	12.6
2005	213	9.32	50	6.07	122	10.1	41	15.8
2006	249	10.2	59	7.02	147	11.3	43	14.5
2007	340	13.3	77	8.51	207	15.4	56	18.4
2008	352	13.6	91	9.48	212	15.9	49	16.4
2009	464	16.5	138	12.3	256	18.5	70	22.9
2010	479	17	155	13.8	270	19.2	54	18.8
2011	544	18.9	172	15.2	323	21.8	49	18
2012	598	20.7	200	18	332	22.5	66	22.1
2013	652	23.8	211	20.2	347	25	94	30.6
2014	673	25	278	24.1	311	25	84	28.9
2015	790	27.3	291	24.7	378	28.1	121	33.1
2016	820	24.7	326	23.5	383	24.4	111	30.7

Table 6: Share of trials receiving NIH funding

	P1 All Trials	P1 Gen. PPM	P1 Rest. PPM	P2 All Trials	P2 Gen. PPM	P2 Rest. PPM	P3 All Trials	P3 Gen. PPM	P3 Rest. PPM
1995	2.10	5.99	3.26	0.00	0.00	0.00	0.00	0.00	0.00
1996	3.54	4.81	3.39	0.00	10.00	11.11	0.00	11.11	11.11
1997	5.56	8.02	3.05	9.09	6.25	0.00	11.11	0.00	0.00
1998	5.34	6.00	1.89	0.00	19.35	6.67	0.00	16.67	0.00
1999	4.08	6.93	2.95	15.38	8.11	0.00	13.04	5.71	0.00
2000	7.29	6.10	2.64	7.41	10.42	10.00	8.00	8.70	5.26
2001	4.73	5.67	3.27	14.63	12.00	0.00	16.67	13.04	0.00
2002	3.91	4.97	2.51	15.22	12.86	3.57	9.76	13.64	3.85
2003	5.13	4.14	2.07	6.67	5.88	2.78	2.70	6.25	2.86
2004	4.08	5.81	1.15	16.18	7.94	2.50	18.33	7.14	2.50
2005	4.34	4.94	1.60	4.48	9.79	3.77	3.45	10.69	3.92
2006	4.05	4.75	0.98	12.16	9.58	1.72	9.38	7.69	1.82
2007	3.22	3.73	0.75	7.34	9.96	0.00	7.87	10.09	0.00
2008	3.29	3.62	0.88	17.24	10.92	4.76	18.27	11.06	5.08
2009	3.54	3.18	0.81	8.43	6.00	0.00	8.11	6.57	0.00
2010	2.74	2.47	0.75	11.35	4.82	1.49	12.73	4.56	1.56
2011	1.96	2.43	0.69	5.14	5.26	0.00	5.85	5.54	0.00
2012	2.07	2.04	0.24	5.63	4.46	0.00	4.78	3.94	0.00
2013	2.66	2.81	0.34	5.84	7.32	0.78	5.63	7.59	0.83
2014	3.13	2.25	0.52	5.04	3.87	0.00	4.90	4.37	0.00
2015	2.63	2.48	0.82	6.52	6.64	0.63	6.52	6.71	0.69
2016	2.67	1.85	0.64	6.47	4.79	0.00	7.23	4.75	0.00
N	43615	57636	30720	2837	4365	1283	2478	3991	1195

Table 7: Predicting PPM trials (linear probability models)

Outcome = PPM trial, generous definition				
	All Years		2005-2016 Only	
Trial start year	0.0038*	0.0038*	0.0050	0.0050
	(0.0014)	(0.0014)	(0.0024)	(0.0024)
Phase 2 Clinical (includes combined phase 2/3 trials)	0.0092	0.0092	0.0116	0.0121
	(0.0095)	(0.0093)	(0.0108)	(0.0106)
Phase 3 Clinical	0.0176	0.0177	0.0200	0.0204
	(0.0148)	(0.0146)	(0.0164)	(0.0162)
U.S. trial site	0.0139***	0.0135**	0.0151***	0.0106**
	(0.0021)	(0.0036)	(0.0024)	(0.0035)
Cancer indication	0.1369***	0.1363***	0.1501***	0.1444***
	(0.0147)	(0.0120)	(0.0183)	(0.0134)
NIH funding	0.0075	0.0074	0.0049	0.0030
	(0.0072)	(0.0067)	(0.0096)	(0.0083)
Biomarker type = genomic	0.2459*	0.2458*	0.2435*	0.2431*
	(0.1111)	(0.1113)	(0.1140)	(0.1142)
U.S. trial * cancer indication		0.0013		0.0134
		(0.0071)		(0.0122)
N	106626	106626	90650	90650
R^2	0.272	0.272	0.280	0.280
Outcome = PPM trial, generous definition				
	All Years		2005-2016 Only	
Trial start year	0.0034*	0.0034*	0.0043	0.0043
	(0.0014)	(0.0014)	(0.0023)	(0.0023)
Phase 2 Clinical (includes combined phase 2/3 trials)	0.0120	0.0121	0.0143	0.0148
	(0.0100)	(0.0099)	(0.0115)	(0.0113)
Phase 3 Clinical	0.0224	0.0226	0.0245	0.0251
	(0.0157)	(0.0155)	(0.0175)	(0.0173)
U.S. trial site	0.0107***	0.0093**	0.0113***	0.0058*
	(0.0017)	(0.0029)	(0.0017)	(0.0024)
Cancer indication	0.1355***	0.1336***	0.1486***	0.1417***
	(0.0140)	(0.0113)	(0.0175)	(0.0127)
NIH funding	0.0050	0.0046	0.0042	0.0018
	(0.0070)	(0.0067)	(0.0086)	(0.0075)
Biomarker type = genomic	0.2210	0.2209	0.2182	0.2177
	(0.1094)	(0.1095)	(0.1120)	(0.1123)
U.S. trial * cancer indication		0.0040		0.0162
		(0.0067)		(0.0116)
N	106626	106626	90650	90650
R^2	0.255	0.255	0.262	0.262

* p<0.05, ** p<0.01, *** p<0.001

All models include a constant; robust standard errors clustered at the level of the ICD-9 chapter

Table 8: Dependent variable: Trial duration in months

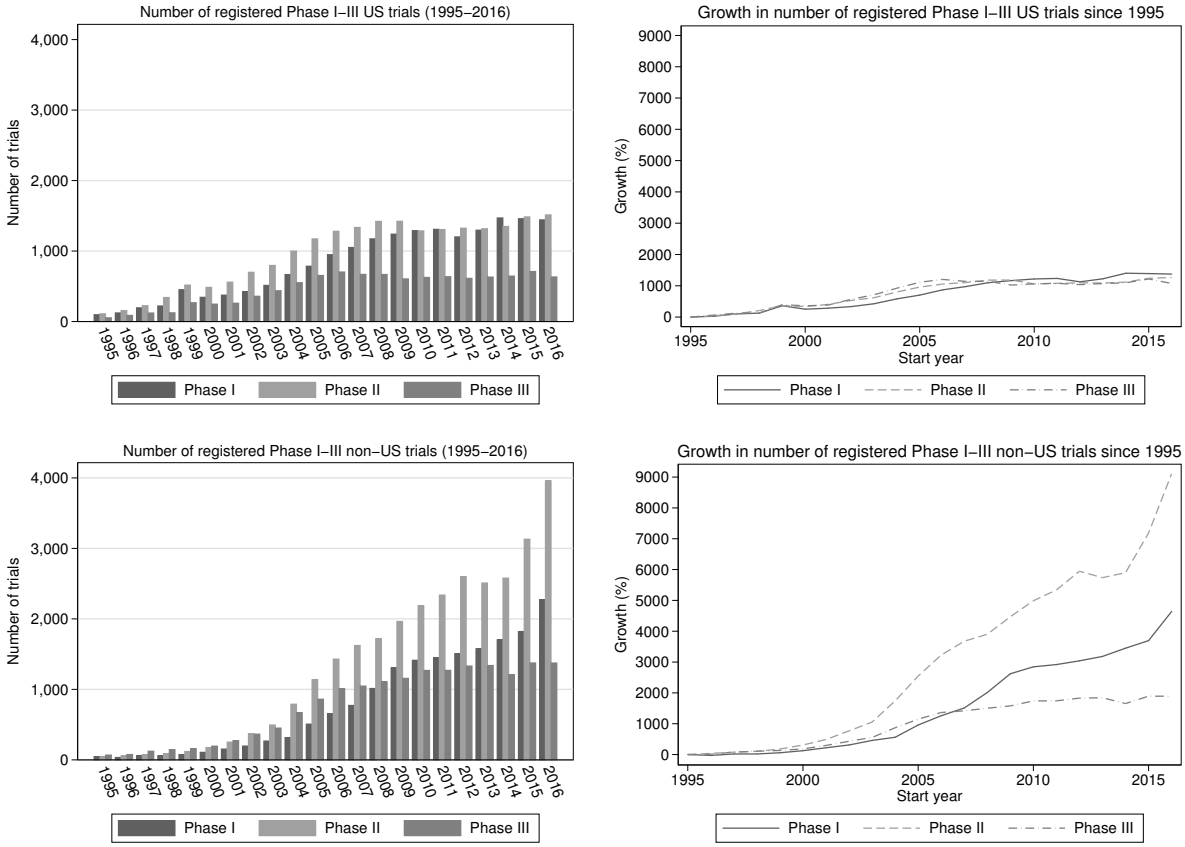
	All Trials	U.S. Trials	U.S. Cancer Trials	
Phase 2 Clinical (includes combined phase 2/3 trials)	6.616***	5.299***	2.246***	2.047***
	(0.220)	(0.332)	(0.574)	(0.605)
Phase 3 Clinical	9.021***	6.132***	9.223***	8.243***
	(0.255)	(0.397)	(1.255)	(1.291)
U.S. trial site	3.564***			
	(0.188)			
Cancer indication	19.146***	17.811***		
	(0.266)	(0.369)		
Received NIH funding	14.237***	13.952***	9.760***	8.668***
	(0.692)	(0.706)	(0.994)	(1.046)
Generous PPM	7.694***	8.078***	8.671***	
	(0.535)	(0.691)	(0.765)	
Biomarker role (detailed): diagnosis				2.900
				(1.490)
Biomarker role (detailed): differential diagnosis				3.968***
				(1.190)
Biomarker role (detailed): predicting drug resistance				-1.919
				(1.289)
Biomarker role (detailed): predicting treatment efficacy				4.813***
				(1.423)
Biomarker role (detailed): predicting treatment toxicity				-0.370
				(1.967)
Biomarker role (detailed): screening				1.078
				(1.591)
Biomarker role (detailed): selection for therapy				2.451*
				(1.215)
Biomarker role (detailed): disease profiling				1.157
				(1.121)
Biomarker role (detailed): monitoring disease progression				-1.518
				(1.502)
Biomarker role (detailed): monitoring treatment efficacy				1.378
				(1.201)
Biomarker role (detailed): monitoring treatment toxicity				-3.168
				(2.547)
Biomarker role (detailed): not determined				-0.004
				(2.527)
Biomarker role (detailed): prognosis				0.346
				(1.374)
Biomarker role (detailed): prognosis - risk stratification				1.110
				(1.757)
Biomarker role (detailed): risk factor				0.164
				(1.154)
Biomarker role (detailed): staging				-4.562***
				(1.324)
Biomarker role (detailed): toxicity profiling				7.606
				(10.024)
N	52929	27861	9315	8443
R^2	0.314	0.286	0.179	0.193

* p<0.05, ** p<0.01, *** p<0.001

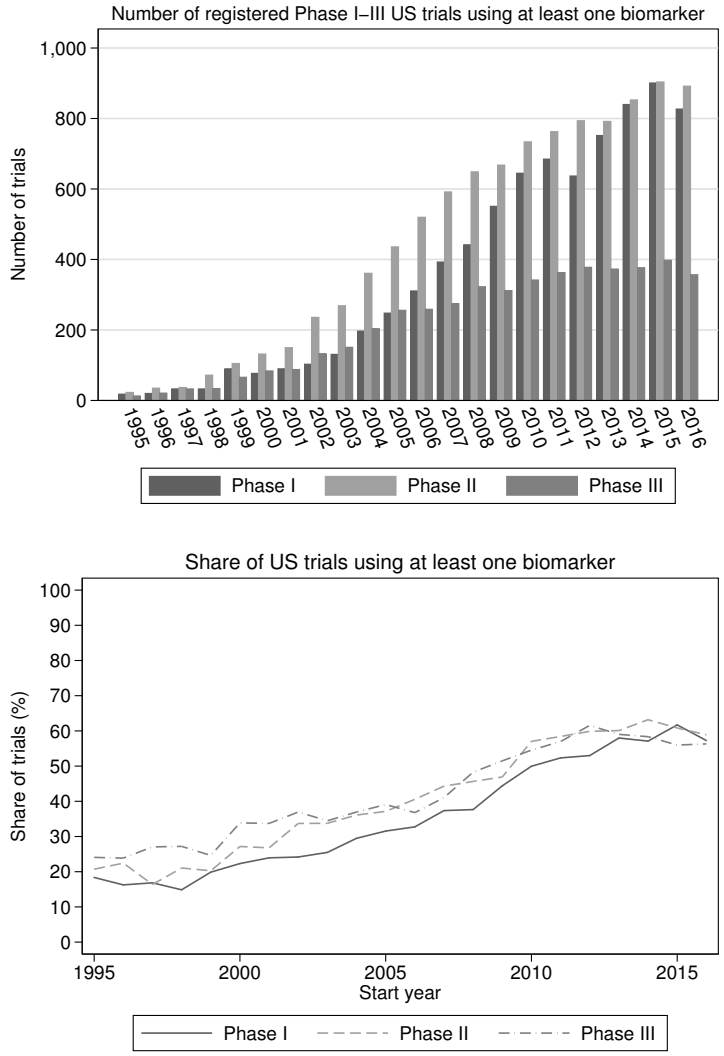
Sample includes all trials launched after 2000 with known end dates. Duration is winsorized to remove extreme outliers. All OLS models include a constant, year fixed effects, and robust standard errors.

Appendices: U.S. Trials

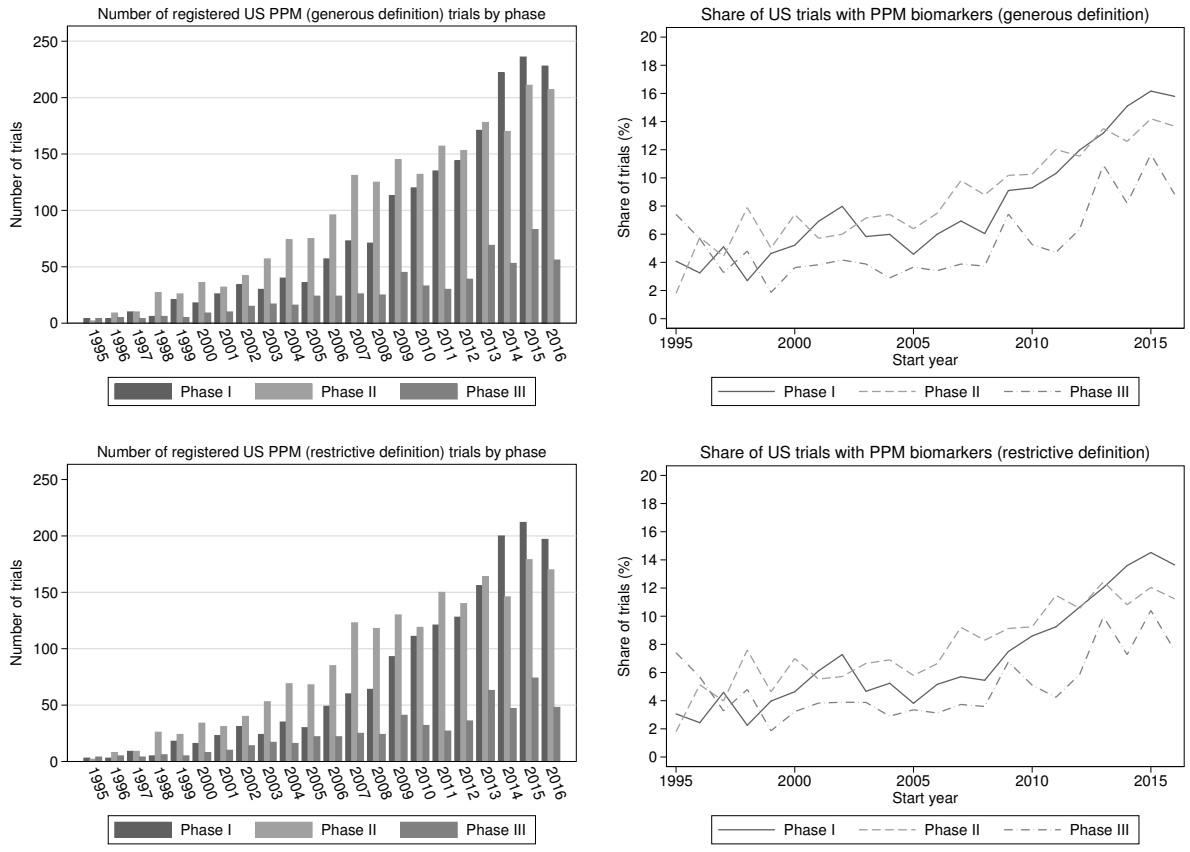
Appendix Figure A: U.S. Clinical trials over time



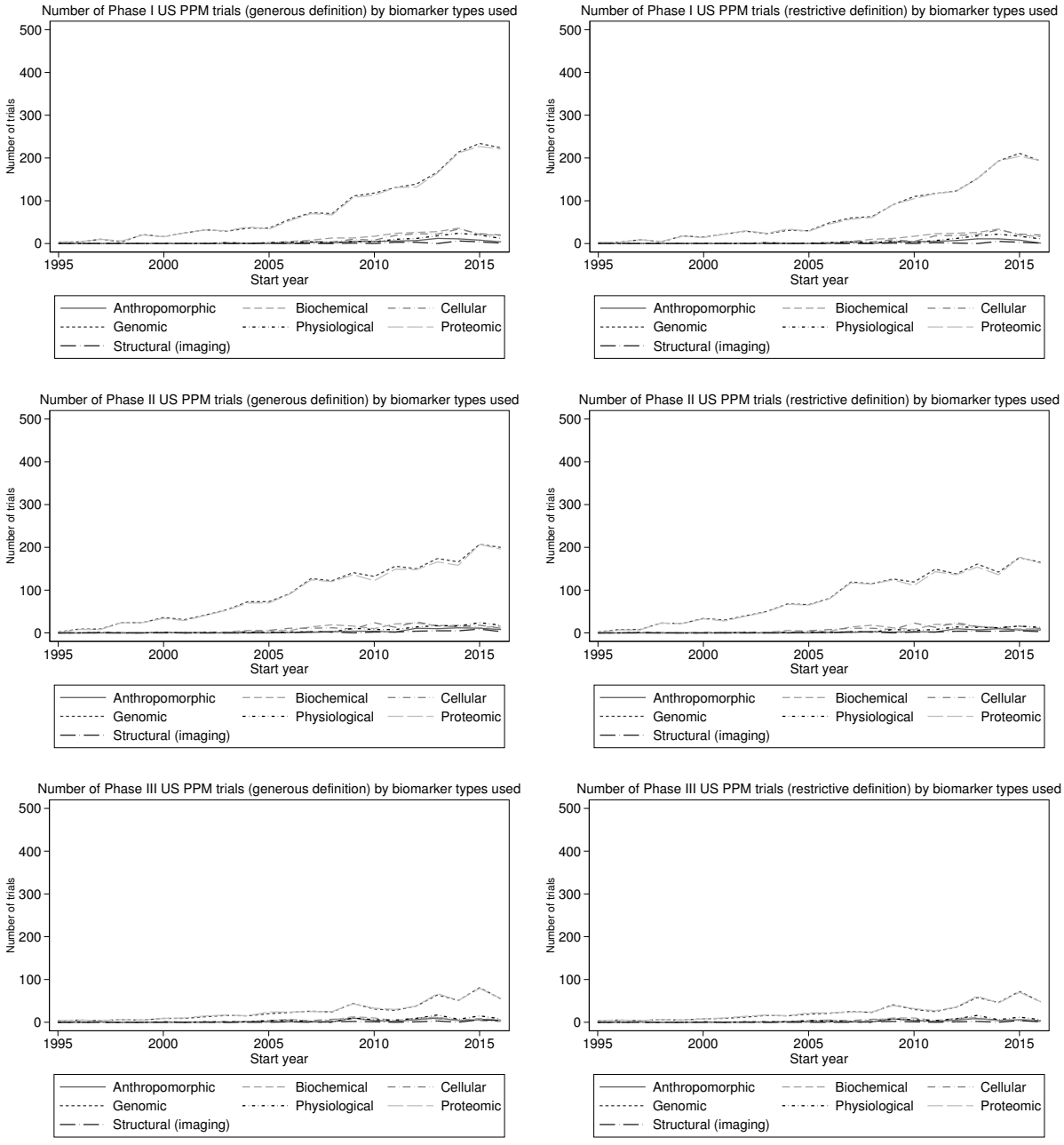
Appendix Figure B: U.S. Clinical trials using biomarkers



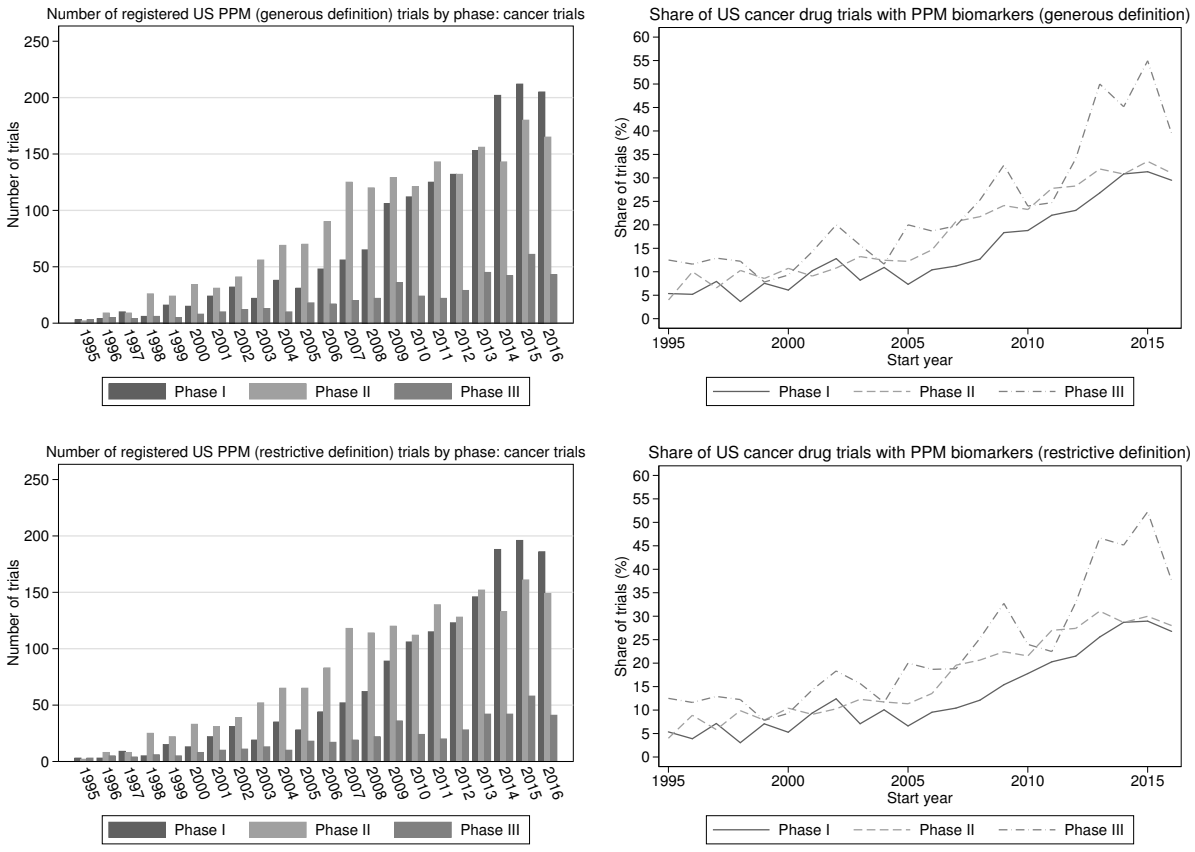
Appendix Figure C: U.S. Clinical trials for PPMs



Appendix Figure D: Types of biomarkers used in U.S. PPM trials



Appendix Figure E: U.S. clinical trials for PPMs, cancer indications only



Appendix Table I: U.S. potential precision medicine (PPM) trials (1995-2016):

	Generous definition							
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	8	6.15	3	5.36	2	4	3	12.5
1996	18	8.57	4	5.19	9	10	5	11.6
1997	23	7.82	10	7.94	9	6.57	4	12.9
1998	38	8.15	6	3.66	26	10.3	6	12.2
1999	45	8.09	16	7.55	24	8.57	5	7.81
2000	57	8.78	15	6.1	34	10.7	8	9.3
2001	65	10.1	24	10.2	31	9.09	10	14.3
2002	85	12.3	32	12.8	41	10.8	12	20
2003	91	11.8	22	8.21	56	13.2	13	15.7
2004	117	11.8	38	10.9	69	12.5	10	11.6
2005	119	11	31	7.33	70	12.2	18	20
2006	155	13.3	48	10.4	90	14.7	17	18.7
2007	201	16.7	56	11.2	125	20.7	20	19.8
2008	207	18	65	12.7	120	21.7	22	25.3
2009	271	22.2	106	18.3	129	24.1	36	32.7
2010	257	21.1	112	18.8	121	23.3	24	24
2011	290	24.8	125	22	143	27.8	22	24.7
2012	293	26.1	132	23.1	132	28.3	29	34.1
2013	354	30.8	153	26.8	156	31.9	45	50
2014	387	31.9	202	30.8	143	30.8	42	45.2
2015	453	34.2	212	31.3	180	33.5	61	55
2016	413	30.9	205	29.5	165	31	43	39.4
	Restrictive definition							
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	8	6.15	3	5.36	2	4	3	12.5
1996	16	7.62	3	3.9	8	8.89	5	11.6
1997	21	7.14	9	7.14	8	5.84	4	12.9
1998	36	7.73	5	3.05	25	9.88	6	12.2
1999	42	7.55	15	7.08	22	7.86	5	7.81
2000	54	8.32	13	5.28	33	10.4	8	9.3
2001	63	9.75	22	9.36	31	9.09	10	14.3
2002	81	11.7	31	12.4	39	10.3	11	18.3
2003	84	10.9	19	7.09	52	12.3	13	15.7
2004	110	11.1	35	10.1	65	11.7	10	11.6
2005	111	10.2	28	6.62	65	11.3	18	20
2006	144	12.4	44	9.54	83	13.5	17	18.7
2007	189	15.7	52	10.4	118	19.6	19	18.8
2008	198	17.2	62	12.1	114	20.7	22	25.3
2009	245	20	89	15.4	120	22.4	36	32.7
2010	242	19.9	106	17.8	112	21.5	24	24
2011	274	23.4	115	20.3	139	27	20	22.5
2012	279	24.8	123	21.5	128	27.4	28	32.9
2013	340	29.6	146	25.6	152	31.1	42	46.7
2014	363	30	188	28.7	133	28.7	42	45.2
2015	415	31.3	196	29	161	30	58	52.3
2016	376	28.1	186	26.8	149	28	41	37.6

Appendix Table II: U.S. potential precision medicine (PPM) trials: cancer only (1995-2016):

Generous definition								
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	8	6.15	3	5.36	2	4	3	12.5
1996	18	8.57	4	5.19	9	10	5	11.6
1997	23	7.82	10	7.94	9	6.57	4	12.9
1998	38	8.15	6	3.66	26	10.3	6	12.2
1999	45	8.09	16	7.55	24	8.57	5	7.81
2000	57	8.78	15	6.1	34	10.7	8	9.3
2001	65	10.1	24	10.2	31	9.09	10	14.3
2002	85	12.3	32	12.8	41	10.8	12	20
2003	91	11.8	22	8.21	56	13.2	13	15.7
2004	117	11.8	38	10.9	69	12.5	10	11.6
2005	119	11	31	7.33	70	12.2	18	20
2006	155	13.3	48	10.4	90	14.7	17	18.7
2007	201	16.7	56	11.2	125	20.7	20	19.8
2008	207	18	65	12.7	120	21.7	22	25.3
2009	271	22.2	106	18.3	129	24.1	36	32.7
2010	257	21.1	112	18.8	121	23.3	24	24
2011	290	24.8	125	22	143	27.8	22	24.7
2012	293	26.1	132	23.1	132	28.3	29	34.1
2013	354	30.8	153	26.8	156	31.9	45	50
2014	387	31.9	202	30.8	143	30.8	42	45.2
2015	453	34.2	212	31.3	180	33.5	61	55
2016	413	30.9	205	29.5	165	31	43	39.4
Restrictive definition								
	All Count	All %	P1 Count	P1 %	P2 Count	P2 %	P3 Count	P3 %
1995	8	6.15	3	5.36	2	4	3	12.5
1996	16	7.62	3	3.9	8	8.89	5	11.6
1997	21	7.14	9	7.14	8	5.84	4	12.9
1998	36	7.73	5	3.05	25	9.88	6	12.2
1999	42	7.55	15	7.08	22	7.86	5	7.81
2000	54	8.32	13	5.28	33	10.4	8	9.3
2001	63	9.75	22	9.36	31	9.09	10	14.3
2002	81	11.7	31	12.4	39	10.3	11	18.3
2003	84	10.9	19	7.09	52	12.3	13	15.7
2004	110	11.1	35	10.1	65	11.7	10	11.6
2005	111	10.2	28	6.62	65	11.3	18	20
2006	144	12.4	44	9.54	83	13.5	17	18.7
2007	189	15.7	52	10.4	118	19.6	19	18.8
2008	198	17.2	62	12.1	114	20.7	22	25.3
2009	245	20	89	15.4	120	22.4	36	32.7
2010	242	19.9	106	17.8	112	21.5	24	24
2011	274	23.4	115	20.3	139	27	20	22.5
2012	279	24.8	123	21.5	128	27.4	28	32.9
2013	340	29.6	146	25.6	152	31.1	42	46.7
2014	363	30	188	28.7	133	28.7	42	45.2
2015	415	31.3	196	29	161	30	58	52.3
2016	376	28.1	186	26.8	149	28	41	37.6

Appendix Table III: Predicting PPM trials (Marginal effects from logit models)

Outcome = PPM trial, generous definition				
	All Years		2005-2016 Only	
Trial start year	0.0029***	0.0029***	0.0045***	0.0045***
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Phase 2 Clinical (includes combined phase 2/3 trials)	0.0119***	0.0114**	0.0144***	0.0138***
	(0.0033)	(0.0036)	(0.0032)	(0.0036)
Phase 3 Clinical	0.0175*	0.0165*	0.0194**	0.0184*
	(0.0068)	(0.0074)	(0.0069)	(0.0075)
U.S. trial site	0.0057	0.0056***	0.0063	0.0062***
	(0.0054)	(0.0016)	(0.0064)	(0.0018)
Cancer indication	0.1207***	0.1206***	0.1294***	0.1293***
	(0.0036)	(0.0031)	(0.0039)	(0.0034)
NIH funding	-0.0061***	-0.0042***	-0.0079***	-0.0056***
	(0.0015)	(0.0006)	(0.0016)	(0.0008)
Biomarker type = genomic	0.2053***	0.2058***	0.2035***	0.2040***
	(0.0067)	(0.0062)	(0.0076)	(0.0069)
N	106626	106626	90650	90650
Outcome = PPM trial, generous definition				
	All Years		2005-2016 Only	
Trial start year	0.0025***	0.0025***	0.0038***	0.0038***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Phase 2 Clinical (includes combined phase 2/3 trials)	0.0139***	0.0137***	0.0162***	0.0158***
	(0.0019)	(0.0022)	(0.0019)	(0.0022)
Phase 3 Clinical	0.0235***	0.0229***	0.0254***	0.0247***
	(0.0036)	(0.0040)	(0.0038)	(0.0043)
U.S. trial site	0.0029	0.0028**	0.0028	0.0027*
	(0.0033)	(0.0009)	(0.0039)	(0.0011)
Cancer indication	0.1202***	0.1201***	0.1288***	0.1288***
	(0.0021)	(0.0017)	(0.0023)	(0.0019)
NIH funding	-0.0060***	-0.0049***	-0.0067***	-0.0054***
	(0.0015)	(0.0007)	(0.0015)	(0.0007)
Biomarker type = genomic	0.1819***	0.1822***	0.1794***	0.1797***
	(0.0041)	(0.0037)	(0.0046)	(0.0041)
N	106626	106626	90650	90650

* p<0.05, ** p<0.01, *** p<0.001

All models include a constant; robust standard errors clustered at the level of the ICD-9 chapter

References:

Bespalov, Anton, et al. "Failed trials for central nervous system disorders do not necessarily invalidate preclinical models and drug targets." *Nature Reviews Drug Discovery* (2016).

Budish, Eric, Benjamin N. Roin, and Heidi Williams. "Do firms underinvest in long-term research? Evidence from cancer clinical trials." The American economic review 105.7 (2015): 2044-2085.

Chandra, A. and Garthwaite, C., 2017. The Economics of Indication-Based Drug Pricing. New England Journal of Medicine, 377(2), pp.103-106.

Costinot, A., Donaldson, D., Kyle, M. and Williams, H., 2016. The more we die, the more we sell? a simple test of the home-market effect (No. w22538). National Bureau of Economic Research.

Cohen, Joshua P., and Abigail E. Felix. "Personalized medicine's bottleneck: diagnostic test evidence and reimbursement." Journal of personalized medicine 4.2 (2014): 163-175.

De Angelis, Catherine, et al. "Clinical trial registration: a statement from the International Committee of Medical Journal Editors." *New England Journal of Medicine* 351.12 (2004): 1250-1251.

Dzau, V.J., Ginsburg, G.S. Realizing the Full Potential of Precision Medicine in Health and Health Care. JAMA. 2016; 316(16):1659–1660.

FDA, Center for Drug Evaluation and Research and Center for Biologics Evaluation and Research. "Guidance for Industry: Information Program on Clinical Trials for Serious or Life-Threatening Diseases and Conditions." (Draft Guidance) *U.S. Department of Health and Human Services, Food and Drug Administration* (2004).

FDA, Center for Drug Evaluation and Research and Center for Biologics Evaluation and Research. "Guidance for Industry E15 Definitions for Genomic Biomarkers, Pharmacogenomics, Pharmacogenetics, Genomic Data and Sample Coding Categories." *U.S. Department of Health and Human Services, Food and Drug Administration* (2008).

Fridlyand, J., Simon, R., Walrath, J., Roach, N., Buller, R., Schenkein, D., Flaherty, K., Allen, J., Sigal, E., & Scher, H. 2013, Considerations for the successful co-development of targeted cancer therapies and companion diagnostics, Nature Reviews Drug Discovery, 12, 10, pp. 743-755

Guedj, I., & Scharfstein, D. (2004). *Organizational scope and investment: Evidence from the drug development strategies and performance of biopharmaceutical firms* (No. w10933). National Bureau of Economic Research.

Hegde, Deepak, and Bhaven Sampat. "Can private money buy public science? Disease group lobbying and federal funding for biomedical research." *Management Science* 61, no. 10 (2015): 2281-2298.

International Trade Administration, Department of Commerce. 2016 ITA Pharmaceuticals Top Markets Report." (2016).

Kanavos, Panos, Alessandra Ferrario, Sotiris Vandoros, and Gerard F. Anderson. "Higher US branded drug prices and spending compared to other countries may stem partly from quick uptake of new drugs." Health affairs 32, no. 4 (2013): 753-761.

Kao, Jennifer L. "R&D Decisions for New Medical Technologies: Evidence from New Use Approvals and Off-Label Uses." *Working Paper* (2017).

Phyo, Aung Pyae, et al. "Antimalarial activity of artefenomel (OZ439), a novel synthetic antimalarial endoperoxide, in patients with Plasmodium falciparum and Plasmodium vivax malaria: an open-label phase 2 trial." *The Lancet Infectious Diseases* 16.1 (2016): 61-69.

Rock Health "2016 Year End Funding Report" Available at https://rockhealth.com/reports/2016-year-end-funding-report-a-reality-check-for-digital-health (2017).

Stern, A.D., Alexander, B.M. and Chandra, A., 2017. How economics can shape precision medicines. Science, 355(6330), pp.1131-1133.

APPENDIX A

This table lists the formal definition of different biomarker types as defined by the FDA-NIH Biomarker Working group (2016)

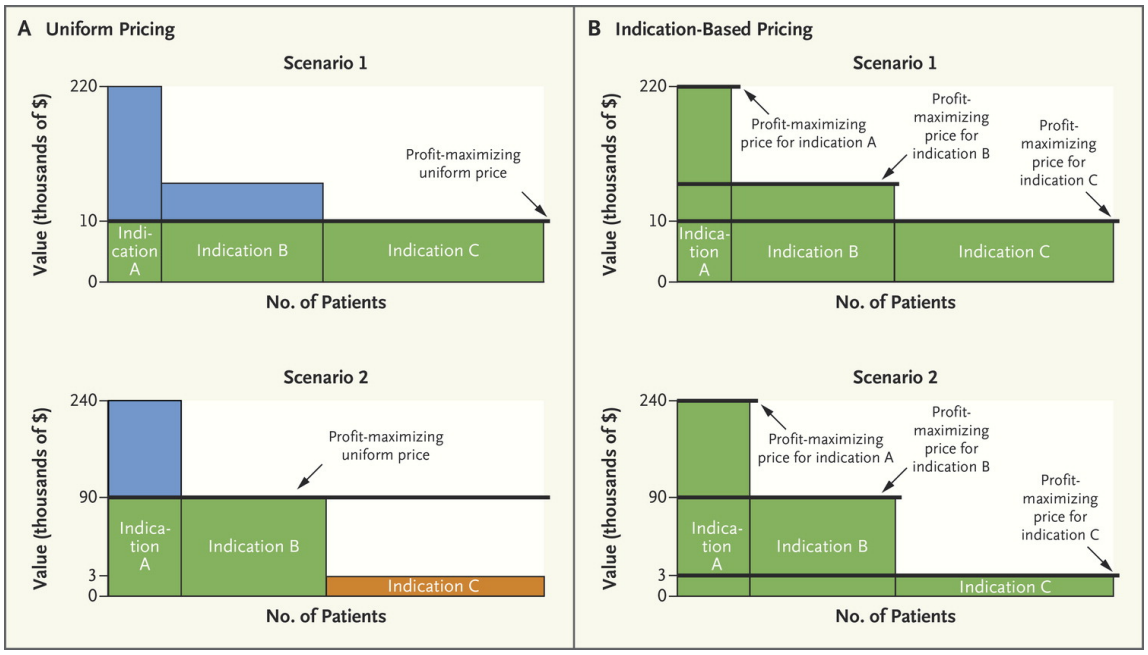
Biomarker type	Official definition	Examples
Diagnostic Biomarker	A biomarker used to detect or confirm presence of a disease or condition of interest or to identify individuals with a subtype of the disease.	1) Sweat chloride may be used as a diagnostic biomarker to confirm cystic fibrosis (Farrell et al. 2008). 2) Glomerular filtration rate (GFR) may be used as a diagnostic biomarker to identify patients with chronic kidney disease (National Kidney Foundation 2002).
Monitoring Biomarker	A biomarker measured serially for assessing status of a disease or medical condition or for evidence of exposure to (or effect of) a medical product or an environmental agent.	1) HIV-RNA may be used as a monitoring biomarker to measure and guide treatment with antiretroviral therapy (ART) (AIDSinfo 2007). 2) Serial measurements of symphysis-fundal height during pregnancy can be used during antenatal screening to detect fetal growth disturbances (Papageorghiou et al. 2016).
Pharmacodynamic / Response Biomarker	A biomarker used to show that a biological response has occurred in an individual who has been exposed to a medical product or an environmental agent.	1) Circulating B lymphocytes may be used as a pharmacodynamic/response biomarker when evaluating patients with systemic lupus erythematosus to assess response to a B-lymphocyte stimulator inhibitor (Stohl and Hilbert 2012). 2) Urinary level of glycosaminoglycans may be used as a pharmacodynamic/response biomarker when evaluating the effect of enzyme replacement therapy for patients with mucopolysaccharidosis type 1 (Jameson et al. 2016).
Predictive Biomarker	A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from exposure to a medical product or an environmental agent.	1) Certain cystic fibrosis transmembrane conductance regulator (CFTR) mutations may be used as predictive biomarkers in clinical trials evaluating treatment for cystic fibrosis, to select patients more likely to respond to particular treatments (Davies et al. 2013). 2) Human leukocyte antigen allele (HLA)–B*5701 genotype may be used as a predictive biomarker to evaluate human immunodeficiency virus (HIV) patients before abacavir treatment, to identify patients at risk for severe skin reactions (AIDSinfo 2007).
Prognostic Biomarker	A biomarker used to identify likelihood of a clinical event, disease recurrence or progression in patients who have the disease or medical condition of interest.	1) Breast Cancer genes 1 and 2 (BRCA1/2) mutations may be used as prognostic biomarkers when evaluating women with breast cancer, to assess the likelihood of a second breast cancer (Basu et al. 2015). 2) Gleason score may be used as a prognostic biomarker when evaluating patients with prostate cancer to assess the likelihood of cancer progression (Epstein et al. 2016; Gordetsky and Epstein 2016).
Safety Biomarker	A biomarker measured before or after an exposure to a medical product or an environmental agent to indicate the likelihood,	1) Hepatic aminotransferases and bilirubin may be used as safety biomarkers when evaluating potential hepatotoxicity (Senior 2014).

	presence, or extent of toxicity as an adverse effect.	2) Serum creatinine may be used as a safety biomarker when evaluating patients on drugs that affect kidney function to monitor for nephrotoxicity (Wasung et al. 2015).
Susceptibility / Risk Biomarker:	A biomarker that indicates the potential for developing a disease or medical condition in an individual who does not currently have clinically apparent disease or the medical condition.	1) Factor V Leiden may be used as a susceptibility/risk biomarker to identify individuals with a predisposition to develop deep vein thrombosis (DVT) (Kujovich 2011). 2) Infection with certain human papillomavirus (HPV) subtypes may be used as a susceptibility/risk biomarker to identify individuals with a predisposition to develop cervical cancer (Khan et al. 2005; Schiffman et al. 2011).

Note: Some examples of biomarkers cited in this appendix may be applicable for more than one type of biomarker. For example, in some cases predictive biomarkers used to identify individuals who are more likely to experience a favorable effect from a drug can also be used as diagnostic biomarkers in the initial detection or confirmation of the disease (e.g. CFTR mutations in Cystic Fibrosis).

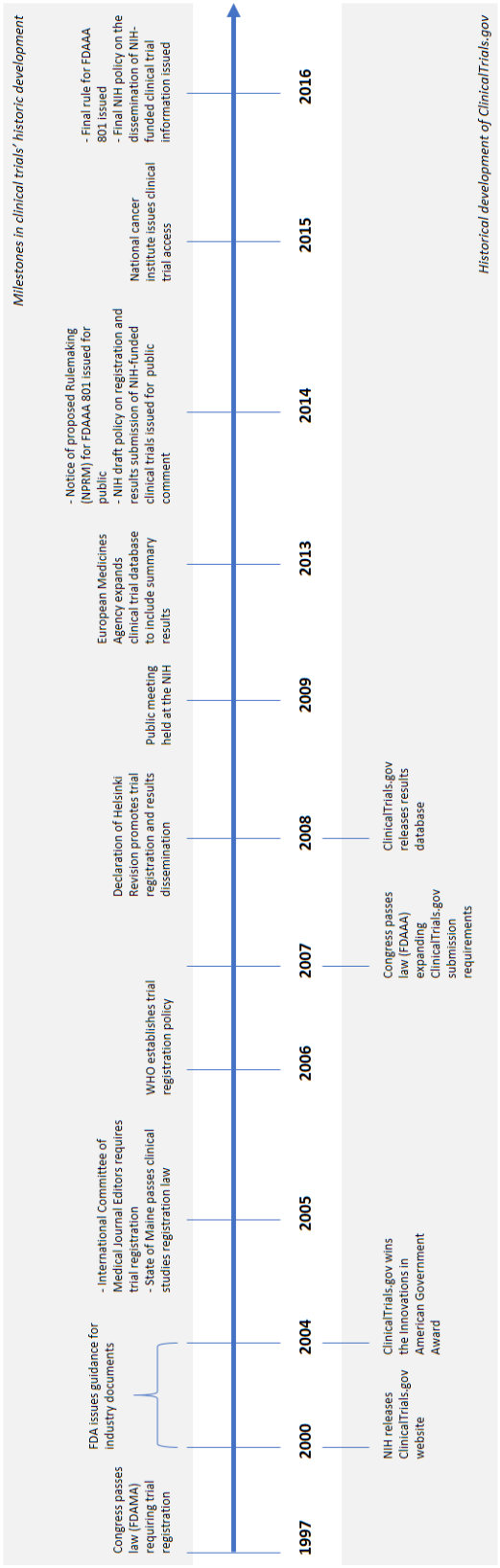
APPENDIX B

Effects of uniform pricing versus indication-based pricing.



In Panel A, the upper graph represents a uniform-pricing context in which patients with indication A receive the most benefit and those with indication C receive the least. The population with indication C is large, and the value of treatment to this group is close to the value for indication B. As a result, the manufacturer's profit-maximizing price allows all patients to obtain the drug. At this price, the manufacturer earns profits represented by the green area. But the firm faces a trade-off. By setting the price in this way, the manufacturer forgoes profits that could be earned by charging higher prices to patients with indications A and B. These forgone profits, represented by the blue areas, are captured by these patients as consumer surplus — the value difference between the most consumers are willing to pay and what they actually pay. The lower graph in Panel A shows a different scenario, in which the product's valuation for patients with indication C is very low. In this case, it's a better trade-off for the manufacturer to set a high price, at which it knows the payer will allow only patients with indications A and B to obtain the drug. The manufacturer accepts the loss of sales to patients with indication C in exchange for higher profits earned from patients with indications A and B. Comparing these graphs, we see that when the valuation of the product for indication C is relatively low, manufacturers set a higher uniform price, the payer curtails sales to pa-tients with indication C (orange area), and patients with indications A and B obtain less consumer surplus than they did in the first scenario

APPENDIX C


Milestones in clinical trials' historic development
Congress passes law (FDAMA) requiring trial registration
FDA issues guidance for industry documents
- International Committee of Medical Journal Editors requires trial registration
- State of Maine passes clinical studies registration law
WHO establishes trial registration policy
Declaration of Helsinki Revision promotes trial registration and results dissemination
Public meeting held at the NIH
European Medicines Agency expands clinical trial database to include summary results
- Notice of proposed Rulemaking (NPRM) for FDAAA 801 issued for public
- NIH draft policy on registration and results submission of NIH-funded clinical trials issued for public comment
National cancer institute issues clinical trial access
- Final rule for FDAAA 801 issued
- Final NIH policy on the dissemination of NIH-funded clinical trial information issued
1997
2000
2004
2005
2006
2007
2008
2009
2013
2014
2015
2016
NIH releases ClinicalTrials.gov website
ClinicalTrials.gov wins the Innovations in American Government Award
Congress passes law (FDAAA) expanding ClinicalTrials.gov submission requirements
ClinicalTrials.gov releases results database
Historical development of ClinicalTrials.gov

Selected Explanation as provided by the Website of ClinicalTrials.gov (2017):

1997: Congress Passes Law (FDAMA) Requiring Trial Registration
The first U.S. Federal law to require trial registration was the Food and Drug Administration Modernization Act of 1997 (FDAMA) (PDF). Section 113 of FDAMA required the National Institutes of Health (NIH) to create a public information resource on certain clinical trials regulated by the Food and Drug Administration (FDA)

2000: NIH Releases ClinicalTrials.gov Web Site
The first version of ClinicalTrials.gov was made available to the public on February 29, 2000. At the time, ClinicalTrials.gov primarily included NIH-funded studies.

2000–2004: FDA Issues Guidance for Industry Documents
In 2000 FDA issued a draft Guidance for Industry document, which provided recommendations for researchers submitting information to ClinicalTrials.gov. A final guidance document that incorporated comments from the public was issued in 2002.

2004: ClinicalTrials.gov Wins the Innovations in American Government Award
The Innovations in American Government Awards program highlights exemplary models of government innovation and advances efforts to address the Nation's most pressing public concerns.

2005: International Committee of Medical Journal Editors Requires Trial Registration
In 2005 the International Committee of Medical Journal Editors (ICMJE) began requiring trial registration as a condition of publication.

2005: State of Maine Passes Clinical Studies Registration Law (Repealed in 2011)
In 2005 the State of Maine passed a law requiring prescription drug manufacturers or labelers to submit clinical study registration and results information to ClinicalTrials.gov. In 2011 the law was repealed; it is no longer in effect.

2006: World Health Organization Establishes Trial Registration Policy
In 2006 the World Health Organization (WHO) stated that all clinical trials should be registered, and it identified a minimum trial registration dataset of 20 items and in 2007 launched the International Clinical Trials Registry Platform (ICTRP).

2007: Congress Passes Law (FDAAA) Expanding ClinicalTrials.gov Submission Requirements
In 2007 the requirements for submission to ClinicalTrials.gov were expanded after Congress passed the Food and Drug Administration Amendments Act of 2007 (FDAAA). Section 801 of FDAAA (FDAAA 801) required more types of trials to be registered; additional trial registration information; and the submission of summary results, including adverse events, for certain trials. The law also included penalties for noncompliance, such as the withholding of NIH grant funding and civil monetary penalties of up to $10,000 a day.

2008: ClinicalTrials.gov Releases Results Database
In September 2008, as required by FDAAA 801, ClinicalTrials.gov began allowing sponsors and principal investigators to submit the results of clinical studies.[26]

[26] The submission of adverse event information was optional when the results database was released but was required beginning in September 2009.

2008: Declaration of Helsinki Revision Promotes Trial Registration and Results Dissemination
In October, 2008 the 59th World Medical Association (WMA) General Assembly amended the Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. Two newly added principles (paragraphs 19 and 30) considered the prospective registration and the public disclosure of study results to be ethical obligations.

2009: Public Meeting Held at the National Institutes of Health
In accordance with FDAAA 801, NIH held a public meeting in April 2009 to solicit input from interested individuals about future regulations that will expand the information on ClinicalTrials.gov.

2013: European Medicines Agency Expands Clinical Trial Database to Include Summary Results
In October 2013 the European Medicines Agency (EMA) released a new version of the European Clinical Trials Database (EudraCT). Notably, the EudraCT summary results data requirements are "substantially aligned" with those of the ClinicalTrials.gov results database.

2014: Notice of Proposed Rulemaking (NPRM) for FDAAA 801 Issued for Public Comment
In November 2014 the U.S. Department of Health and Human Services issued a notice of proposed rulemaking (NPRM) describing the proposed requirements and procedures for registering and submitting the results, including adverse events, of clinical trials on ClinicalTrials.gov, in accordance with FDAAA 801.

2014: NIH Draft Policy on Registration and Results Submission of NIH-Funded Clinical Trials Issued for Public Comment.
In November 2014 NIH proposed a policy to ensure that every clinical trial (see the Revised NIH Definition of "Clinical Trial") that receives NIH funding is registered on ClinicalTrials.gov and has summary results submitted and posted in a timely manner, whether subject to FDAAA 801 or not.

2015: National Cancer Institute Issues Clinical Trial Access Policy
In January, 2015 the NIH National Cancer Institute (NCI) issued its Policy Ensuring Public Availability of Results from NCI-supported Clinical Trials. The policy states, "Final Trial Results are expected to be reported in a publicly accessible manner within twelve (12) months of the Trial's Primary Completion Date regardless of whether the clinical trial was completed as planned or terminated earlier."

2016: Final Rule for FDAAA 801 Issued
In September 2016, the U.S. Department of Health and Human Services issued a Final Rule for Clinical Trials Registration and Results Information Submission (42 CFR Part 11) that clarifies and expands the regulatory requirements and procedures for submitting registration and summary results information of clinical trials on ClinicalTrials.gov, in accordance with FDAAA 801. The final rule is intended to make it clear to sponsors, investigators, and the public which trials must be submitted, when they must be submitted, and whether compliance has been achieved.

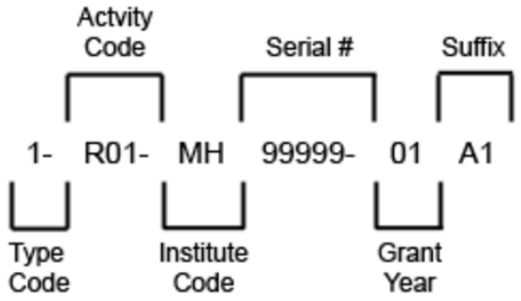
2016: Final NIH Policy on the Dissemination of NIH-funded Clinical Trial Information Issued
In September 2016, NIH issued a final policy to promote broad and responsible dissemination of information from NIH-funded clinical trials through ClinicalTrials.gov. Under this policy, every clinical trial funded in whole or in part by NIH is expected to be registered on ClinicalTrials.gov and have summary results information submitted and posted in a timely manner, whether subject to FDAAA 801 or not.

APPENDIX D

place-holder

APPENDIX E

The following explanation of NIH grant numbers are provided by the NIMH website (2017):
The parts of a complete NIH grant number indicate the following: type, activity code, Institute, serial #, grant year, and (possibly) a suffix. For example, the grant number: 1-R01-MH99999-01A1 indicates:



1- This is the **Type Code**. The most common types are:

- 1- never previously funded grants — that is, a new/first time grant application.
- 2- competing continuations — that is, a grant application that was previously funded for a period of time. This new continuing period of support requires peer review.
- 5- non-competing continuations — that is, a grant application that has been funded and is in the midst of its support period. For each year of the support period awarded, there is an administrative review of progress before the next annual installment of support is issued (no peer review is needed). The application that the PI submits as part of this process is called a "non-competing continuation application," and it contains a "progress report" for the period of support just completed.

R01- **Activity Code** indicates the type of grant mechanism. Examples include R01s (investigator initiated research grant), R03s (small grants), R13s (conference support grants), "K"s (career awards), "T"s (institutional training awards), etc.

MH- **Institute Code** identifies the NIH Institute with primary responsibility for payment of this application. For example, MH = National Institute of Mental Health (NIMH) and DA=National Institute on Drug Abuse (NIDA). Each NIH Institute has a two-letter code associated with it.

99999- **Serial Number** provides a unique identification to the project and is assigned sequentially for newly submitted applications. The Serial Number remains the same for as long as a project is active, even when the PI submits a competing continuation for a new period of support.

01- **Grant Year**. "01" indicates the first year of a grant application or funded grant.

A1- **Suffix**. "A1" indicates that the application was submitted once previously but did not receive a sufficiently strong priority score to merit funding. This application is an amended version of the original one also called a "resubmission." At NIH, an R01 may be submitted up to three separate times for review (i.e., an A2 application is the last amended version permitted). Other suffix terms are also used. For example, "S1" refers to a competing supplement request for a currently funded project.

7. APPENDIX SOURCES:

AIDSinfo. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. December 1, 2007. Accessed October 2016. Available at: https://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv-guidelines/7/hla-b--5701-screening

Basu NN, Ingham S, Hodson J, Lalloo F, Bulman M, Howell A, Evans DG. Risk of contralateral breast cancer in BRCA1 and BRCA2 mutation carriers: a 30-year semi-prospective analysis. Fam Cancer. 2015 Dec;14(4):531–8. doi: 10.1007/s10689-015-9825-9. PubMed PMID: 26239694.

ClinicalTrials.gov. History, Policies, and Laws. July, 2017. Accessed August 2017. Available at: https://clinicaltrials.gov/ct2/about-site/history

Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. Grading Committee. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. Am J Surg Pathol. 2016 Feb;40(2):244–52. doi: 10.1097/PAS.0000000000000530. PubMed PMID: 26492179.

Farrell PM, Rosenstein BJ, White TB, Accurso FJ, Castellani C, Cutting GR, Durie PR, Legrys VA, Massie J, Parad RB, Rock MJ, Campbell PW 3rd. Cystic Fibrosis Foundation. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. J Pediatr. 2008 Aug;153(2):S4–S14. doi: 10.1016/j.jpeds.2008.05.005. PubMed PMID: 18639722.

FDA-NIH Biomarker Working Group. "BEST (Biomarkers, EndpointS, and other Tools) Resource." (2016).

Gordetsky J, Epstein J. Grading of prostatic adenocarcinoma: current state and prognostic implications. Diagn Pathol. 2016 Mar 9;11:25. doi: 10.1186/s13000-016-0478-2. PubMed PMID: 26956509.

Jameson E, Jones S, Remmington T. Enzyme replacement therapy with laronidase (Aldurazyme®) for treating mucopolysaccharidosis type I. Cochrane Database Syst Rev. 2016 Apr 1;4:CD009354. doi: 10.1002/14651858.CD009354.pub4. PubMed PMID: 27033167.

Khan MJ, Castle PE, Lorincz AT, Wacholder S, Sherman M, Scott DR, Rush BB, Glass AG, Schiffman M. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. J Natl Cancer Inst. 2005 Jul 20;97(14):1072–9. doi: 10.1093/jnci/dji187. PubMed PMID: 16030305.

Kujovich JL, Factor V. Leiden thrombophilia. Genet Med. 2011 Jan;13(1):1–16. doi: 10.1097/GIM.0b013e3181faa0f2. PubMed PMID: 21116184.

National Kidney Foundation. K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease: Evaluation, Classification and Stratification. Am J Kidney Dis 39:S1-S266, 2002 (suppl 1). Accessed December 2016. Available at: https://www.kidney.org/sites/default/files/docs/ckd_evaluation_classification_stratification.pdf

"NIMH » Research Funding Frequently Asked Questions (FAQs).". https://www.nimh.nih.gov/funding/grant-writing-and-application-process/research-funding-frequently-asked-questions-faqs.shtml (accessed August 8, 2017).

Papageorghiou A, Ohuma E, Gravett M, Hirst J, Silveira M, Lambert A, Carvalho M, Jaffer Y, Altman D, Noble J, Bertino E, Purwar M, Pang R, Ismail L, Victora C, Bhutta Z, Kennedy S, Villar J. International standards for symphysis-fundal height based on serial measurements from the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project: prospective cohort study in eight countries. BMJ. 2016 Oct;355:i5662. doi: 10.1136/bmj.i5662. PubMed PMID: 27821614.

Schiffman M, Glass AG, Wentzensen N, Rush BB, Castle PE, Scott DR, Buckland J, Sherman ME, Rydzak G, Kirk P, Lorincz AT, Wacholder S, Burk RD. A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study. Cancer Epidemiol Biomarkers Prev. 2011 Jul;20(7):1398–409. doi: 10.1158/1055-9965.EPI-11-0206. PubMed PMID: 21602310.

Senior JR. Evolution of the Food and Drug Administration approach to liver safety assessment for new drugs: current status and challenges. Drug Saf. 2014 Nov;37 Suppl 1:S9–17. doi: 10.1007/s40264-014-0182-7. PubMed PMID: 25352324.

Stohl W, Hilbert DM. The discovery and development of belimumab: the anti-BLyS-lupus connection. Nat Biotechnol. 2012 Jan 9;30(1):69–77. doi: 10.1038/nbt.2076. PubMed PMID: 22231104.

Wasung ME, Chawla LS, Madero M. Biomarkers of renal function, which and when? Clin Chim Acta. 2015 Jan 1;438:350–7. doi: 10.1016/j.cca.2014.08.039. PubMed PMID: 25195004.