# Representation and Extrapolation: Evidence from Clinical Trials

Marcella Alsan*

Maya Durvasula†

Harsh Gupta†

Joshua Schwartzstein‡

Heidi Williams§¶

October 7, 2022

**Abstract**

This article examines the consequences and causes of low enrollment of Black patients in clinical trials. We develop a simple model of similarity-based extrapolation that predicts that evidence is more relevant for decision-making by physicians and patients when it is more representative of the group that is being treated. This generates the key result that the perceived benefit of a medicine for a group depends not only on the average benefit from a trial, but also on the share of patients from that group who were enrolled in the trial. In survey experiments, we find that physicians who care for Black patients are more willing to prescribe drugs tested in representative samples, an effect substantial enough to close observed gaps in the prescribing rates of new medicines. Black patients update more on drug efficacy when the sample that the drug is tested on is more representative, reducing Black-White patient gaps in beliefs about whether the drug will work as described. Despite these benefits of representative data, our framework predicts that those who have benefited more from past medical breakthroughs are less costly to enroll in the present, leading to persistence in who is represented in the evidence base.

*Harvard Kennedy School and NBER.

†Stanford University.

‡Harvard Business School.

§Stanford University and NBER.

¶Contact: marcella_alsan@hks.harvard.edu, maya.durvasula@stanford.edu, hgupta13@stanford.edu, jschwartzstein@hbs.edu, hlwill@stanford.edu. We are grateful to Alyce Adams, Anna Aizer, Michelle Andrasik, Arthur Applbaum, David Autor, John Beshears, Sally Bock, Emily Breza, Amitabh Chandra, Niteesh Choudhry, Katherine Coffman, David Costanzo, Amy Finkelstein, Christopher Hucks-Ortiz, Damon Jones, Rem Koning, Lisa Larrimore Ouellette, Corinne Low, Elisa Macchi, Kyle Myers, Nathan Nunn, Gautam Rao, Andrei Shleifer, Stefanie Stantcheva, Adi Sunderam, and seminar participants from MIT, Stanford, Beth Israel Deaconess Hospital, Massachusetts General Hospital, Brookings Institute, Warwick, Stockholm School of Economics, Rutgers, San Diego State, Georgia State, Vanderbilt, UCSD, Yale, University of Chicago, ICSSI 2022, EAAMO 2022, and the Federal Food and Drug Administration, for helpful comments. Nick Shankar, Lukas Leister, Anne Fogarty, Emma Ronzetti, Xingyou Ye, and Zahra Thabet provided exceptional research assistance. We thank Harlan Krumholz, Joseph Ross, John Welsh and Research!America for sharing data and Michele Andrasik, Sally Bock, David Costanzo, and Christopher Hucks-Ortiz from the Fred Hutchinson Cancer Center for sharing experiences and expertise. The study is approved by the IRB at Harvard University (IRB21-1384 and IRB22-0017) and registered at the AEA RCT Registry (AEARCTR-0008957 and AEARCTR-0008959). We gratefully acknowledge financial support from the National Science Foundation under Grant No. DGE-1656518, the Knight Hennessy Scholars Program and the MacArthur Foundation.

*As a physician caring for patients in an urban safety-net setting and wanting to provide the best evidence-based preventive care... I would spend as much time on the science as I devoted to reinforcing with patients why they should still trust these guidelines and the process, despite the unrepresentative populations in the evidence base.*

– Dr. Kirsten Bibbins-Domingo, Editor-in-Chief, *Journal of the American Medical Association* (National Academies of Sciences, Engineering, and Medicine Report 2022)

# I Introduction

Innovation does not benefit everyone equally (Aghion et al. 2019; Jones and Kim 2018; Kline et al. 2019). Research investments skew towards developing technologies appropriate for more profitable groups (Cutler, Meara and Richards-Shubik 2012; Jaravel 2019; Kremer and Glennerster 2004; Michelman and Msall 2021), and diffusion often occurs faster among the well-connected or well-educated (Agha and Molitor 2018; Foster and Rosenzweig 2010; Glied and Lleras-Muney 2008; Hamilton et al. 2021; Papageorge 2016; Skinner and Staiger 2005, 2015). In this article, we explore a third dimension of innovation and inequality. We ask whether the low enrollment of certain groups in the R&D process (Koning, Samila and Ferguson 2021) creates gaps in how much group members use those technologies. Put differently, does *how* a technology is developed affect *who* adopts it?

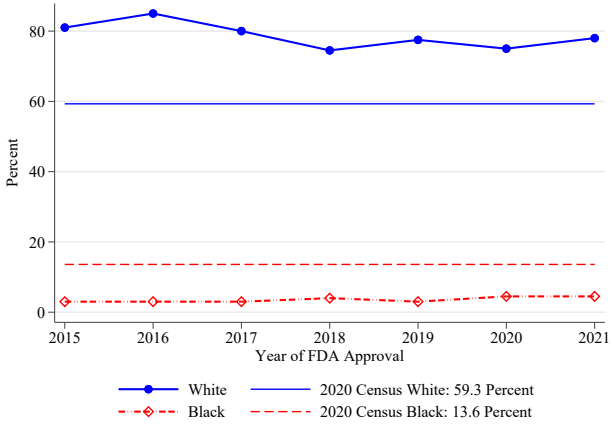
Our context is new drug approval in the United States, where information on drug safety and efficacy — generated from clinical trials on human subjects — must be submitted to the U.S. Food and Drug Administration (FDA) before the drug can be sold. Racial inequality in both the production of clinical evidence and the eventual diffusion of products is commonplace (Ding and Glied 2022; Elhussein et al. 2022*a*,*b*; Jung and Feldman 2017; McCoy et al. 2019; Wang et al. 2007). As Figure I documents, Black patients are consistently underrepresented in clinical trials relative to their share in the U.S. population (Panels a and b) and are similarly underrepresented in prescriptions for newly approved medications (Panels c and d).[1] [2] The median clinical trial includes only five percent Black participants — less than half the population share of Black Americans.[3]

[1]Representation generally requires a benchmark. Throughout this article, we use the two pragmatic and most commonly used benchmarks in the literature: disease burden and population share. Census population is the implied metric in this plot, but we note that Black patients are often even more underrepresented relative to their disease burden (Green et al. 2022). Black underrepresentation has been stable over a longer time series than plotted in Panel (a) (Appendix Figure B1) and across ClinicalTrials.gov and FDA Drug Trials Snapshots data sets (Appendix Figure B2). In contrast, the participation of female patients — another group that has been historically excluded from clinical trials — has been increasing over time and in recent trials is comparable to the female population (see Appendix Figure B3). However, there are several conditions where women remain underrepresented in trials (Gupta 2022; Sosinsky et al. 2022; Feldman et al. 2019; Steinberg et al. 2021). There are other demographic groups (*e.g.*, Hispanic) that are underrepresented in medical research. We focus on Black Americans, in particular, for several reasons, including the long history of (involuntary) medical research in Black communities in the U.S. and low life expectancy (Arias et al. 2022). See Section VI.1 for additional detail.

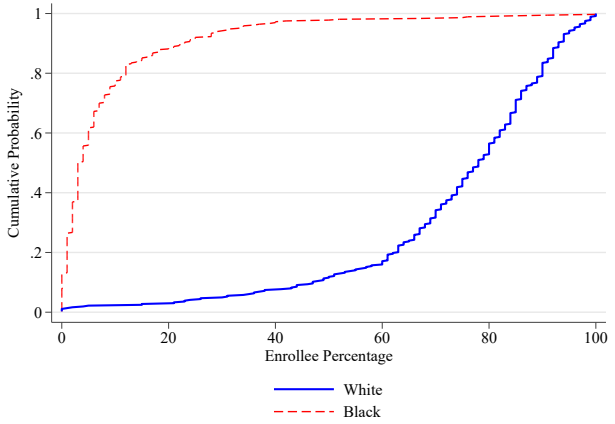
[2]We plot prescribing rates of new drugs per 1000 individuals in each racial group in Appendix Figure B4.

[3]The median pivotal trial for a newly FDA-approved drug in 2021 included 4.5 percent Black participants while the median pivotal trial for a newly FDA-approved drug between 2015–2021 included 3.0 percent Black participants (see Figure I Panel a).

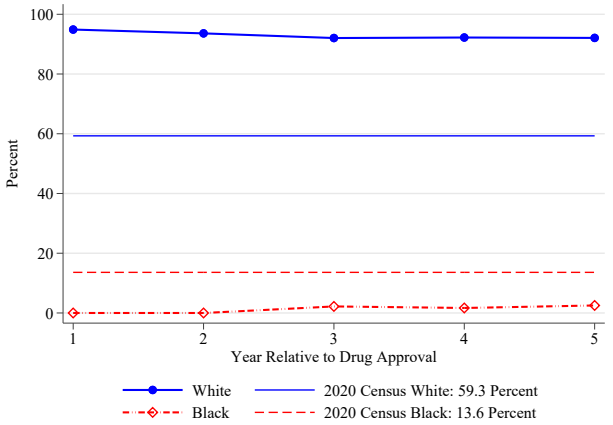
Figure I: Racial Inequality in the Development and Distribution of New Drugs



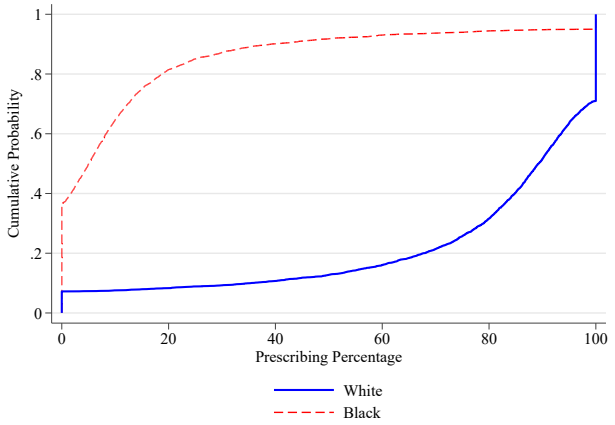

(a) Clinical Trials Participation




(b) CDF of Trial Enrollees




(c) Prescriptions of New Drugs




(d) CDF of New Drug Prescribing

*Notes:* Panel (a) plots the median enrollee percentage by race (Black and White) for pivotal clinical trials, studies that support new drug applications to the FDA, over time. Panel (b) plots the cumulative probability of enrollee percentage for pivotal clinical trials by race. Panel (c) plots the median new drug prescription percentage by race in each year relative to its approval. Panel (d) plots the cumulative probability of new drug prescription percentage by race. In all figures, red dotted lines report shares for Black individuals and blue solid lines report shares for White individuals. Dashed lines in Panels (a) and (c) plot population shares by race in the U.S. as reported in the 2020 Census (Black population share is 13.6 percent and non-Hispanic White population share is 59.3 percent; U.S. Census Bureau 2021). Data for Panels (a) and (b) are drawn from the FDA Drug Trials Snapshots data, and data for Panels (c) and (d) are from the Medical Expenditure Panel Survey. See Appendix H.1.2 and H.2 for details on data.

Firms and regulators have faced pressure to address racial diversity in clinical trials in recent years, yet its complex causes have posed challenges for policymakers. These challenges include differential access to health care and academic research centers (Chandra and Skinner 2003; George, Duran and Norris 2014). They also include mistrust of medical research grounded in contemporaneous (*e.g.*, Roberts 2012; Hoffman et al. 2016) and historical (*e.g.*, Alsan and Wanamaker 2018; Eli, Logan and Miloucheva 2019) discrimination by the health care system and more widely (see Darity, Mullen and Slaughter 2022; Logan and Myers Jr. 2022).

Although gaps in trial enrollment are well-documented, the consequences, if any, have not been rigorously studied. Two natural questions emerge: First, does representative data matter to physicians and patients? Second, if so, why are such data not (endogenously) supplied by the market? To address the first question, we conduct two survey experiments designed to understand physician and patient reactions to trial evidence. To address the second question, we turn to a theoretical framework that sheds light on how underrepresentation may persist, even if representative data leads to higher drug demand, and identifies potential levers for policy intervention, which we then assess in the context of case studies.

Our framework models how physicians and patients interpret the evidence that supports new technologies when making decisions about whether to adopt them. Through their instruction in evidence-based medicine (EBM), physicians are trained to consider whether a new product would work similarly well in their patients as those in its trial. Inspired by this process and the role reasoning by similarity and analogy plays in belief formation (*e.g.*, Gilboa and Schmeidler 1995; Mullainathan, Schwartzstein and Shleifer 2008; Bordalo, Gennaioli and Shleifer 2020; Bordalo et al. 2022; Malmendier and Veldkamp 2022), we develop a model of similarity-based extrapolation. We assume that people update more readily from evidence when their patients (in the case of doctors) or people like them (in the case of patients) have more in common with the experimental sample. Our framework incorporates this assumption in a simple way: It assumes doctors and their patients have in mind a model where a given group characteristic (*e.g.*, race) could be correlated with drug efficacy and update model parameters using Bayes' rule. A key result of our framework is that — conditional on trial data — the perceived benefit of a drug will be increasing not only in the average reported efficacy, but also increasing at a decreasing rate in the share of one's own group in the trial.

These predictions imply that a profit-maximizing firm could increase sales by recruiting a more representative sample. However, the trade-off in doing so is cost — our framework suggests that a history of underrepresentation in (voluntary) research leads Black patients to anticipate lower benefits of trial enrollment, making recruitment more costly. With the status quo recruitment infrastructure, firms enroll the lowest cost individuals — perpetuating doubt about whether trial findings extrapolate to other groups and generating a cycle of underrepresentation.

To empirically assess whether representation affects clinical decisions and health behavior, we designed and conducted two survey experiments among patients and physicians. After completing a short module

eliciting patient panel characteristics, physicians viewed profiles of diabetes drugs, including the drug's mechanism of action and the design of the supporting clinical trials. For each profile, the share of Black trial subjects and average drug efficacy in trials were cross-randomized from distributions of values collected in a comprehensive search of clinical literature. In order to have sufficient variation in both sample demographics and efficacy within the mechanism of action of a given drug, the drugs shown were hypothetical — but were based on recently developed drugs to treat diabetes.[4] After viewing each profile, physicians were asked to indicate their intent to prescribe the drug to patients in their care.

A separate experiment was designed for patients since they must fill and adhere to a prescription in order for health gains to be realized. We recruited 275 patients with diagnosed hypertension who identified as either White or Black. We then assessed their interest in a novel therapy to treat hypertension that had been tested in two sites with varying shares of Black participants. Other product characteristics, including drug efficacy in lowering blood pressure, were held constant.

We find that physicians are more willing to prescribe drugs tested on representative samples. A one standard deviation increase in the share of Black trial participants increases physician prescribing intention for a given drug by 0.11 standard deviation units. The magnitude of this effect on prescribing is medically meaningful and equivalent to roughly half the standardized effect of the drug's efficacy. It also correlates strongly with donation behavior to campaigns aimed at boosting trial participation for underrepresented minority communities measured a few weeks after the initial intervention. In pre-specified heterogeneity analyses, we find that the effect of increasing Black representation in a clinical trial sample on prescribing intention is close to zero for doctors who do not routinely see Black patients and rises steeply in the share of a physician's own patients who are Black.

Similarly, in our patient experiment, when Black respondents were presented with a representative trial, they viewed the drug in question as significantly more relevant for their own blood pressure control and were 20 percentage points more likely to state that the drug will work as well for them as it was shown to work in the trial. In contrast and consistent with the model's prediction of diminishing returns to representation, we do not find significant effects associated with trial composition for White patients. The combination of physician and patient results suggest doctors are broadly acting as agents for their patients.

Survey experiments are important tools for uncovering peoples' mental models and values (Stantcheva 2022), but are subject to critiques, such as experimenter demand and social desirability bias. Our experiments were designed to mitigate such concerns. First, we used neutral recruitment materials stating our aim was to understand views on medical research.[5] Second, we recruited both White and Black patients. If the response to sample representation was solely due to social desirability, we might expect to find similar effects for both groups (we do not). Third, survey responses correlate with actual

[4]We informed physicians that the drugs were hypothetical so they would not try to prescribe them after the experiment.
[5]Only 11.5 percent of physicians and 7.1 percent of patients attrited after consent and this was not differential across arms.

donation behavior in a follow-up study.

A related concern is that our experiment may have informed patients and doctors about something that they did not already know about — *i.e.*, the composition of clinical trials. If so, our results might overstate the degree to which trial representation influences treatment choices. To better understand baseline knowledge in our study populations, we reviewed literature on how doctors evaluate trials and obtained data on patients' knowledge regarding medical research (see Section G). Physicians trained at accredited medical colleges in the U.S. are explicitly taught to consider the applicability of trial findings to their own patients. A typical question from EBM training is: "Are the participants in the study *similar* enough to my patient?" (Masic, Miokovic and Muhamedagic 2008).[6]

In our survey, 72 percent of physicians reported that they have been asked by patients whether a new medicine will "work in people like me." Data from the non-profit Research!America reveal that Black and White respondents are aware of clinical trials (80 percent and 88 percent, respectively). However, Black respondents are less likely to believe science benefits them and less likely to consent if invited to participate in clinical trials than White respondents. Two additional pieces of qualitative evidence suggest trial representation is taken into account by (at least some) doctors and patients: one comes from stakeholder quotes compiled in the writing of a recent National Academy of Medicine (NASEM 2022) report and another comes from the association between more representative clinical trials and higher prescribing rates for new drugs among Black patients (see Section VI.1 and Appendix Table D2).

Turning to mechanisms, we find that doctors — and to a greater extent, patients — lack confidence in extrapolating from samples that are not representative of them or their patients. This is true of both Black patients (when extrapolating across racial groups) and White patients (when extrapolating across geographies). One question is whether this hesitancy to extrapolate, especially among doctors, is a mistake. It's a difficult question to answer for several reasons. First, precisely because representation is so low, clinical trials offer limited direct evidence on this question. The best evidence we could find on the frequency of racial heterogeneity in pivotal trials is from Green et al. (2022), who show that out of 290 new drug approvals in the FDA Drug Trials Snapshots data, approximately 80 percent did not report treatment effects for Black patients separately. Among those that did, 91.4 percent and 98.1 percent found no difference in benefits and side effects, respectively. Second, because the mechanism of action of medications (*i.e.*, a drug's pharmacodynamics) are often incompletely specified, it is difficult to provide assurances that the findings will extrapolate across patients with different characteristics without trial evidence. Third, there is a strong relationship between social class and race in the U.S. that could affect pharmacokinetics, or how the drug is metabolized. Indeed, in our experiments, respondents cited the possibility of biological, socioeconomic, and environmental differences that could alter drug performance as rationales for their lack of confidence. Fourth, even if physicians believe findings do extrapolate, they might internalize patients' lack of confidence for a variety of reasons (Ellis and McGuire 1986), including that it might impact patient adherence. Our qualitative findings from doctors

[6]Emphasis added.

explaining why they care about representation include concerns regarding treatment effect heterogeneity and concern for patients' views.

Importantly, we find increasing the representativeness of medical research can reduce health gradients. Physicians treating Black patients are considerably less willing to prescribe drugs approved on the basis of unrepresentative trials — at all levels of drug efficacy — as compared to physicians who treat White patients, mirroring the racial prescription gap observed in the Medical Expenditure Panel Survey (see Panels c and d of Figure I). When clinical trial samples are more representative of Black patients, however, this gap disappears. The difference between the share of Black and White patients who believe that the drug will work as well for them as it did in clinical trials is also eliminated when respondents are shown results generated from more representative data. These findings suggest that policies that increase representation in the evidence base for new technologies could narrow gaps in their adoption.[7]

Although such policies may take many forms, we discuss case studies of successful investments in what we call inclusive infrastructure, which our framework suggests could break the cycle of underrepresentation. We document considerable variation in trial representation across diseases and contrast two especially different cases: cancer and HIV/AIDS. Although research into both diseases is supported by large, coordinated networks that share many resources, Black patients are well-represented in HIV/AIDS trials and poorly represented in cancer trials, relative to both population share and disease burden benchmarks. To understand the origins of these differences, we draw on interviews with clinical trials networks, qualitative research, and administrative data. We highlight two key features that differentiated HIV/AIDS trials: engagement with priority population communities from protocol design to recruitment, and site selection in and around safety net hospitals. These differences may explain both its more representative evidence base and, more suggestively, its higher diffusion rates of new products.[8]

## Related Literature

Our work contributes to a growing literature that seeks to understand the role of innovation in creating or exacerbating health inequality. Previous studies have focused on how endogenous (demand-pull) investment can affect the composition of resulting technologies. Most closely related is Cutler, Meara and Richards-Shubik (2012), who find that allocation of NIH grant funding disproportionately flows towards majority groups when physicians "treat what they see," widening health gradients in settings

[7]We find a reduction in health disparities, despite not providing information on subgroup-specific treatment effects. As our model makes clear, if people start with a model that places positive probability on race mattering for treatment effects, then this is consistent with Bayesian reasoning (or similarity-based extrapolation more broadly): seeing a greater share of one's own race in the trial makes it more likely that the trial results will apply to them. This suggests trials that represent the target population are beneficial for reducing health disparities, even when they are insufficiently powered to detect racial differences.

[8]Our discussion regarding the location of clinical trial study sites is related to the issue of site selection bias defined in Allcott (2015) as well as to a larger literature on the positive selection of study participants. In the context of clinical trials, Basu and Gujral (2020) model how selection creates a wedge between average treatment effects in the trial and marginal treatment effects for non-participants (*i.e.*, efficacy vs. effectiveness). Positive selection is also important when computing cost-effectiveness for certain preventive health recommendations, see Einav et al. (2020), Kowalski (2022), and Oster (2020).

where disease burden differs across groups. Michelman and Msall (2021) highlight the harm from regulatory restrictions on female participation in early-stage clinical trials, which dampens patent activity for female-specific conditions. Other scholarship focuses attention on how product characteristics affect diffusion. Papageorge (2016) develops a dynamic structural model of demand for medical treatment when patients trade-off health and work experience, illustrating how side effects associated with HIV medication could affect treatment decisions among employed persons. Hamilton et al. (2021) extend this model, describing more generally how patient preferences exert a demand externality, tilting innovation towards less efficacious drugs and lowering overall experimentation.

We build on these important contributions by developing and testing an alternative link between innovation and inequality: we ask whether unequal representation in the R&D process can induce inequality directly by making it more difficult for people to extrapolate from the data to their situation.

Our project is facilitated by the growing use of survey experiments in economics, which allows researchers to unpack reasoning and uncover viewpoints on pressing issues that might otherwise be difficult to credibly observe (Bordalo et al. 2016, 2019, 2022; Alesina, Miano and Stantcheva 2022; Elías, Lacetera and Macis 2019; Haaland, Roth and Wohlfart Forthcoming; Kuziemko et al. 2015; Kesselheim et al. 2012; Stantcheva 2022). It also connects to research seeking to understand how end-users of evidence, such as policymakers (*e.g.*, Hjort et al. 2021) or (in our case) physicians and patients, value different facets of the data-generating process.

The remainder of this paper proceeds as follows. Section II provides background information on clinical trials and relevant history. In Section III, we formalize how representative clinical trials may matter to patients and physicians. Section IV describes our two experiments. Section V presents our experimental results. We conclude by drawing lessons from case studies of successful efforts to improve representation in medical research.

# II Background

This section discusses the institutional context of clinical research, including trial financing and costs, the regulatory review process, and factors that shape enrollment. We also describe how doctors and patients learn about new drugs and trial results. The features highlighted below are then incorporated into our framework. Appendix G provides additional details.

## II.1 Clinical Trials Landscape

### II.1.1 The Drug Development Process

Before a new drug may be marketed in the United States, the FDA must deem it to be both safe and effective. Sponsors seeking to obtain FDA approval typically conduct clinical trials — randomized

evaluations of the new drug relative to a placebo or current standard of care (National Institutes of Health 2017). Data drawn from ClinicalTrials.gov, the largest global registry of clinical trials, suggest that private firms are the most frequent single primary sponsor of clinical trials (42 percent), an order of magnitude more frequent than U.S. federal agencies (4 percent).[9] The remainder of clinical trials are sponsored by academic institutions, hospitals, and non-profit organizations.[10]

The drug approval process begins when sponsors identify a promising lead compound — the core component of what will become a drug. Sponsors typically file initial patent applications on the drug just prior to beginning Phase I clinical trials.[11] When firms begin clinical testing, they also file investigational new drug (IND) applications, which draw on data from pre-clinical testing. Patent terms are twenty years long, though firms may receive other forms of market exclusivity that can extend effective patent life.

Drug sponsors must complete three stages of clinical testing before applying for marketing approval. Phase I trials are intended to establish safety, determine appropriate dosages, and identify side effects. Phase II and III trials test efficacy, monitor safety, and compare the product to existing alternatives. Whereas Phase I trials often recruit a small number of healthy volunteers, Phase II and III trials recruit from the target patient population and may enroll thousands of people. Drug approval hinges on so-called "pivotal" trials, which are typically Phase III trials that aim to demonstrate efficacy.

### II.1.2 The Cost of Clinical Trials

Clinical research is expensive. Recent estimates suggest that the median cost of a pivotal clinical trial providing evidence of efficacy to the FDA is about $19 million (Moore et al. 2018).[12] Industry reports suggest the most expensive step of the clinical trial process is the recruitment of patient participants in Phases II and III (Sertkaya et al. 2014). Accrual rates — the speed with which a trial can recruit eligible patients — are cited as the most common reason for trial delays and, in some cases, failure. Slower accrual rates can lengthen clinical trial periods and erode patent life (Budish, Roin and Williams 2015).[13] Thus, trial sponsors aim to identify and enroll patients as quickly as possible, often contracting with third parties that specialize in clinical trial enrollment, and sometimes moving operations overseas

[9] See also Ehrhardt, Appel and Meinert (2015) for evidence of relative importance of industry sponsorship. We collected data on clinical trials that study products approved for sale in the United States, which are regulated by U.S. agencies. See Appendix H.1.1 for details on data.

[10] These institutions are flagged as "Other" in ClinicalTrials.gov. We reviewed institutions in this set to confirm that our interpretation of "Other" was correct.

[11] We verify this using data drawn from the U.S. Federal Register. In nearly all cases, core patents are filed just before the beginning of clinical testing. See Budish, Roin and Williams (2015) for a discussion on the timing of initial patent filing.

[12] This estimate reported in Moore et al. (2018) draws on proprietary data and estimates the costs of pivotal trials associated with new drugs approved by the FDA in 2015 and 2016. Note that both smaller and larger estimates of trial cost have been reported in academic literature. For example, DiMasia, Grabowski and Hansen (2016) estimated the median cost of a Phase III trial as $200 million.

[13] Long accrual periods are akin to "commercialization lags" discussed in Budish, Roin and Williams (2015): to the extent that long periods of recruitment reduce the profitability of a trial, firms have strong incentives to avoid such cases.

where recruitment costs tend to be lower (Qiao, Alexander and Moore 2019).

The cost — in terms of both money and time — of enrolling a new patient in a trial also varies across demographic groups. To the best of our knowledge, there are no publicly available estimates of trial recruitment costs.[14] Two proxies do, however, provide some quantitative information on the size of these cost differences. First, consider the cost to Moderna of improving representation in one of the highest-stakes clinical trials in recent history: its first-generation SARS-CoV-2 vaccine. In September 2020, the company announced that trial enrollment was going to be slowed for the explicit purpose of improving representation of patients from racial and ethnic minorities in the trial. Moderna's stock price fell eight percent upon the announcement (Appendix Figure B5) (Tirrell and Miller 2020). A second illustrative data point is the cost of recruiting Black experimental subjects in a lower-cost environment: online marketing surveys. In Appendix Figure B6, we plot price quotes for U.S.-based respondents that we received for our own study from three large survey firms. All three firms quoted higher prices to recruit Black respondents as compared to White respondents — with prices ranging from 4 to 130 percent more to recruit a Black respondent. We endogenize these cost differences and explore their effects in our conceptual framework (Section III).

### II.1.3 Enrollment Patterns and Barriers to Participation

The higher cost of recruiting Black Americans may play a role in explaining the trial enrollment patterns described in Figure I. Black Americans comprise just five percent of trial enrollees in the median clinical trial — far less than the 13.6 percent of the U.S. population that they comprise (U.S. Census Bureau 2021). This level has remained flat since data collection efforts began (Appendix Figure B1). Based on the Research!America survey data, Black Americans are less likely to have confidence in research institutions, to believe science is beneficial for them or to enroll in clinical trials (Table I).[15] These findings mirror those of our own survey data: an analysis of open-text responses reveals that Black patients are more likely to cite trust, privacy and racism as reasons not to enroll whereas White patients cite logistical barriers and co-morbidities (Appendix Figure B7).

### II.1.4 Clinical Trials Data

Upon successful completion of the three phases of clinical trials, sponsors submit new drug applications (NDAs) to the FDA. Based on these data, the FDA determines whether the drug will be approved for sale in the U.S. and for which specific indications. Currently, the FDA only requires that a drug is proven efficacious for the "target population," which in practice translates into patients with the targeted condition. Most trials are therefore powered to detect a mean difference in the primary endpoint

[14] For one discussion, see Barel, Bomen, and Morten (2021).

[15] Note that these gaps are relatively constant when we control for income, education, and political affiliation (see Appendix Table C1). Though we also note that conditioning on many characteristics may not always be appropriate when quantifying racial gaps (see Appendix A.1).

Table I: Views on Science and Clinical Trials Among U.S. Respondents

| | Black Respondents (1) | White Respondents (2) | Difference (3) |
|---|---|---|---|
| Confidence in Research Institutions | 2.829 | 3.082 | -0.253*** |
| | (0.963) | (0.822) | |
| Heard of Clinical Trial | 0.796 | 0.875 | -0.079*** |
| | (0.374) | (0.339) | |
| Would Enroll in Clinical Trial if Doctor Recommends | 0.783 | 0.837 | -0.054*** |
| | (0.384) | (0.379) | |
| Trust Not Reason for Lack of Enrollment | 0.432 | 0.536 | -0.104*** |
| | (0.463) | (0.514) | |
| Science is Beneficial | 0.284 | 0.383 | -0.099*** |
| | (0.419) | (0.493) | |
| Would Get FDA-Approved Vaccine | 2.907 | 3.069 | -0.163 |
| | (1.024) | (1.099) | |
| Kling-Liebman-Katz (KLK) Index | 0.926 | 1.152 | -0.226*** |
| | (0.406) | (0.521) | |

*Notes:* Table reports the survey responses among Black and White U.S.-based individuals across a number of questions regarding science. Data are drawn from a nationally representative survey conducted by Research!America in the years 2013, 2017, and 2021 and pooled. *Heard of Clinical Trial*, *Trust*, and *Science is Beneficial* are dichotomous variables. Other variables are on an ordinal scale. See Appendix H.3.5 for details on data. Standard deviations are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

between treatment and control groups, and not to detect subgroup-specific treatment effects, which are uncommonly reported (Green et al. 2022). The most common statistic reported in abstracts and quoted in advertisements is therefore a drug's average treatment effect, as demonstrated in the trial. Demographic characteristics of the sample are typically provided in the first table (the balance table) of journal articles or in the short description of the study population in drug advertisements.[16]

### II.1.5 The Market for New Drugs

Although analogous approval processes occur worldwide, approval in the U.S. market is critical for pharmaceutical firms since it has an outsized role in the industry: U.S. sales were projected to account for nearly 50 percent of the $1.2 trillion in global pharmaceutical revenues earned in 2020 (IQVIA 2015) and a disproportionate share of pharmaceutical net income (Goldman and Lakdawalla 2018; Ledley et al. 2020). The U.S. is an important market since it currently lacks price controls other countries use to curtail spending and is permissive with respect to marketing. Given these features of the market, the framework in Section III focuses on firms choosing recruitment strategies to maximize profit, and we

[16]Sample size and measures of statistical significance and precision are also reported in abstracts. We reviewed publications associated with ∼500 clinical trials, including 341 referenced in Welsh et al. (2018) and ∼150 trials associated with products approved for sale in the United States, published between 2015 and 2020. In nearly all cases, average effects of interventions were reported in the abstracts. Nearly all trials included some demographic information in a balance table, and approximately 50 percent reported race.

model demand for new drugs from the perspective of consumers situated within the U.S.

## II.2 Demand for New Drugs in the U.S.

### II.2.1 How Physicians Learn about New Drugs

RCTs are considered the gold standard for causal inference in medicine and have been since their popularization by the British Medical Research Council and subsequent adoption by the FDA in 1962 (Cochrane 1972). EBM is a step-by-step process that facilitates the "reasonable use of modern best evidence in making decisions about the care of individual patients" (Martí-Carvajal 2020, p.∼1). EBM's five steps aim to integrate clinical experience, patient values and research findings (Blanco et al. 2014).[17]

After physicians complete their formal training, trial information is often accessed via multiple sources including ClinicalTrials.gov, academic journals, society or national practice guidelines, pharmaceutical representatives, medical conferences, and, more informally, via online and in-person social networks.[18] To maintain an active medical license, family practice and internal medicine physicians must participate in continuing medical education (CME) as required by their state board on an annual basis (Federation of State Medical Boards 2022). In addition, the American Board of Internal Medicine requires physicians to earn 100 "maintenance of certification" points each year and pass an exam (American Board of Internal Medicine 2022). Similar guidelines exist for the American Academy of Family Physicians (American Academy of Family Physicians 2022). In addition to meeting requirements set by state and federal professional associations, doctors have other reasons for keeping up to date with the literature, including the desire to help their patients (Doximity 2014).

### II.2.2 How Patients Learn about New Drugs

Patients learn about new drugs through their physicians and via advertisements. The U.S. is one of two countries that allows firms to market medications directly to patients.[19] Between 2016 and 2018, firms spent $17.8 billion on direct-to-consumer advertising associated with 553 unique drugs (U.S. Government Accountability Office 2021). Patient advocacy groups in the United States are also key in disseminating information about new drugs — lists of trials and summaries of evidence exist for nearly

[17]The steps include: a) problem definition; b) search for wanted sources of information; c) critical evaluation of the information; d) application of information to the patient; and e) efficacy evaluation of this application on patient. It is in this penultimate step — application of the information to the particular patient — that the specific question is asked: "Are the participants in the study similar enough to my patients?" (Masic, Miokovic and Muhamedagic 2008, p.∼222).

[18]ClinicalTrials.gov is widely used as a source of information. As of April 2019, the website received more than 215 million page views per month and 145,000 unique visitors daily. See their website for additional details. Between 2014 and 2018, drug and medical device companies spent roughly $2 billion on physician payments, including speaking and consulting, meals, and travel (Ornstein, Weber and Jones 2019).

[19]The U.S. and New Zealand are the only two countries that allow direct to consumers advertising (DTCA) that includes explicit claims about the product (Schwartz and Woloshin 2019). This targeting can be very precise based on search history and sometimes includes links to clinical information.

all major categories of disease. Data from Research!America show that 80 percent of Black respondents and 88 percent of White respondents had heard of clinical trials (Table I).

Even if patients are aware of trials, they may not know precise details about their composition. However, individuals who have been systematically excluded may nevertheless intuit such patterns, contributing to the more pessimistic views on science and research observed in Table I. We document in our survey of primary care physicians that 72 percent report having been asked by their patients about whether a new medication will "work in people like me." The share of physicians asked this question on a regular basis is higher among those that treat Black patients (Appendix Figure B8). Our theoretical framework considers beliefs and behavior of U.S.-based patient-physician dyads with access to information on average treatment effects and demographics from trials — we then report results from experimentally manipulating these two features of trials in Section IV.

# III Organizing Framework

The framework presented below serves three purposes. First, it formalizes how representation in the trial process affects perceived benefits of new drugs for patients and their doctors, yielding predictions we can then test experimentally. Second, it deepens understanding of why underrepresentation of Black patients is an equilibrium outcome, requiring us to move beyond experimental predictions and model the costs and benefits to firms conducting clinical trials. Third, it clarifies why patterns of exclusion have been so persistent, identifying an intertemporal externality associated with historical exclusion.

## III.1 Physicians and Patients

Physicians and patients use clinical trial information to understand the benefits of a new treatment and participation in clinical research. Both agents are important end-users of clinical trial information: physicians are the gatekeepers of prescriptions, whereas patients' adherence behavior determines whether prescribed drugs will have the intended salubrious effect. To abstract from strategic interactions between physicians and patients and instead focus on the core issues surrounding consequences and causes of low representation, we make two assumptions that guarantee that a doctor's decision of whether to prescribe a treatment (or recommend trial participation) aligns with a patient's decision to adhere to the prescription (or participate in the trial). First, we follow the standard assumption that everyone shares a common prior. Second, we assume doctors are agents for patients and share their objective function.[20]

[20]These assumptions simplify the presentation of the model, but it will be clear that the intuitions that arise from the model do not hinge on them.

### III.1.1 Physician and Patient Beliefs

The assessments of patient-doctor dyad $i$ are influenced by the current and historical trial data, which are in turn influenced by the current firm's recruitment strategy and historic recruitment strategies of all firms (see Section III.2). Suppose the benefits to treatment for patient $i$ equal $b_i \in \{0, \tilde{b}\}$ for $\tilde{b} > 0$, where benefits are measured relative to not getting treatment. That is, the treatment either doesn't work ($b_i = 0$) or works ($b_i = \tilde{b}$), and $\tilde{b}$ parameterizes the stakes of the disease-treatment combination. The likelihood that the treatment works for a patient with characteristics $x_i$ is given by $\theta(x_i) \equiv \Pr(b_i = \tilde{b}|x_i) \in [0,1]$. Overall, then, the perceived benefit of treatment, $\hat{b}_i$, is:

$$\hat{b}_i = \tilde{b} \times \mathbb{E}_i\left[\theta(x_i) \mid \text{trial data}\right],$$

where $\mathbb{E}_i[\cdot]$ is the expectation of dyad $i$ on whether the treatment will work and this expectation is conditioned on data available at the time of the decision. The assumption that everyone applies the same model of inference allows us to simplify the presentation of the model in two ways. First, the expectation operator is identical across all dyads and we can write $\mathbb{E}_i[\cdot]$ as $\mathbb{E}[\cdot]$. Second, the perceived benefit of treatment $\hat{b}_i$ only depends on $i$ through $i$'s characteristics $x_i$ (*i.e.*, it is not heterogeneous conditional on $x_i$), so whenever it does not cause confusion we will write $\hat{b}_i$ as a function of $x_i$ and the available data $h$: $\hat{b}_i = \hat{b}(x_i; h)$.

To focus attention on racial inequality and simplify the exposition, assume $x_i$ is uni-dimensional and in $\{0,1\}$, where $x_i = 0$ corresponds to "White" and $x_i = 1$ to "Black". As noted above, clinical trials rarely report subgroup analyses. Instead, data from a given trial $t \in \{1,\ldots,T\}$ consist of the combination of the average reported efficacy and fraction of Black participants, $(\bar{b}_t, \bar{x}_t)$. Average efficacy is defined as $\bar{b}_t \equiv \tilde{b}_t \times k_t/N_t$, where $\tilde{b}_t$ denotes the benefits of the treatment if successful, $k_t$ the number of trial participants for whom the treatment was in fact successful, and $N_t$ the number of trial participants.[21] The fraction of Black trial participants simply equals $\sum_j x_j/N_t$, where the summation is taken over the trial participants. The complete history of trial data $h$ equals $h^{T-1} = (\bar{b}_t, \bar{x}_t)_{t=1}^{T-1}$ before treatment $t = T$'s trial is run and equals $h^T = (\bar{b}_t, \bar{x}_t)_{t=1}^{T}$ after. Our focus will be on beliefs about this treatment $t = T$ and, when it does not cause confusion, will omit the $t$ subscript when referring to it.

The *key* assumption underlying patients' and doctors' model of inference $\hat{b}(\cdot)$ is that, in assessing the likelihood of treatment success for patients with characteristics $x_i$, they extrapolate more from data on patients with those characteristics than from data on patients with different characteristics. For patients, this could reflect learning from similarity, central to a wide variety of evidence-backed frameworks in psychology and economics.[22] For doctors, this is consistent with evidence-based medicine (see

[21] For simplicity, we abstract from the need for a control group and also assume $\tilde{b}_t$ is known to the firm ahead of the trial, while $k_t$ is stochastic and revealed by the trial.

[22] Such learning includes case-based learning (Gilboa and Schmeidler 1995), analogical reasoning (Jehiel 2005; Mullainathan, Schwartzstein and Shleifer 2008), associative learning (Bordalo, Gennaioli and Shleifer 2020; Mullainathan 2002), reinforcement learning (Daw 2014), and the idea that information from similar sources "resonates" more than information from dissimilar sources (Malmendier and Veldkamp 2022).

Section II.2.1). Formally, people form beliefs about $\theta(x_i)$ and hence $\hat{b}(x_i;h)$ by attaching probability $m$ to characteristic $x_i$ *m*attering.[23] We then have

$$\hat{b}(x_i;h) = m \times \left(\tilde{b} \times \mathbb{E}[\theta(x_i)|h, x_i \text{ matters}]\right) + (1-m) \times \left(\tilde{b} \times \mathbb{E}[\theta(x_i)|h, x_i \text{ doesn't matter}]\right).$$

To generate simple closed-form expressions for the above expectations, we assume priors over $\theta$ are in the Beta family. If $\theta(x_i)$ is distributed according to Beta distributions prior to the trial data for treatment $T$, with parameters $(\alpha(x_i;h^{T-1}), \beta(x_i;h^{T-1}))$ conditional on $x_i$ mattering and parameters $(\alpha(h^{T-1}), \beta(h^{T-1}))$ conditional on $x_i$ not mattering, then:

$$\hat{b}(x_i;h^{T-1}) = m \times \underbrace{\left(\tilde{b} \times \frac{\alpha(x_i;h^{T-1})}{\alpha(x_i;h^{T-1}) + \beta(x_i;h^{T-1})}\right)}_{\text{posterior estimate of } \hat{b} \text{ conditional on } x_i \text{ mattering}} + (1-m) \times \underbrace{\left(\tilde{b} \times \frac{\alpha(h^{T-1})}{\alpha(h^{T-1}) + \beta(h^{T-1})}\right)}_{\text{posterior estimate of } \hat{b} \text{ conditional on } x_i \text{ not mattering}} \quad (1)$$

We set initial conditions for these parameters such that $\alpha(x_i,h^0) = \beta(x_i,h^0) = \alpha(h^0) = \beta(h^0)$ (*i.e.*, in the absence of trial data agents assess the likelihood of treatment success as 0.5).

If clinical trial data are available, people form priors on the efficacy of novel treatments under investigation (more on this below), and update their beliefs once trial data on those treatments become available. We assume people attribute fraction $\bar{x}_T(x_i)$ of the overall number $k_T$ of successes reported in the trial to study participants with $x_i$, where $\bar{x}_T(x_i)$ equals the fraction of trial participants with characteristics $x_i$.[24]

Given this assumption, they then update their beliefs from trial data on treatment $T$ according to Bayesian updating (see Appendix F.1 for precise equations).[25]

[23] In the case that $x_i$ matters, they believe $\theta(x_i = 0)$ is statistically independent of $\theta(x_i = 1)$, so evidence on whether the treatment works on people with $x_i = 0$ does not speak to whether it works on people with $x_i = 1$ and vice-versa. In the case that $x_i$ doesn't matter, they believe $\theta(x_i = 0)$ equals $\theta(x_i = 1)$. We simplify by assuming that $m$ is fixed over time — *i.e.*, that people don't update their beliefs about $m$.

[24] Recall, the FDA does not require (and trials are therefore not powered to report) treatment efficacy conditional on $x_i$. The assumption that successes attributable to participants with $x_i$ scale with their proportion in the trial is a conservative assumption on how people "fill in" missing data as it rules out physician- or patient-assumed heterogeneous trial efficacy as the mechanism deriving our predictions. Relaxing this assumption would increase the importance of representation in our model.

[25] As is standard, people end up placing some weight on the prior (given by $\alpha/(\alpha+\beta)$) and some on the empirical success probability in the trial ($k/N$).

**Proposition 1.** *Supposing $m > 0$ is fixed and average trial efficacy ($\frac{k_T}{N_T}$) exceeds prior-belief ratios $\left(\frac{\alpha(x_i;h^{T-1})}{\alpha(x_i;h^{T-1})+\beta(x_i;h^{T-1})}$ and $\frac{\alpha(h^{T-1})}{\alpha(h^{T-1})+\beta(h^{T-1})}\right)$, then:*

1. $\frac{\partial \hat{b}(x_i;h^T)}{\partial k_T} > 0$*: the perceived benefit of a treatment to a patient is increasing in efficacy, as measured within the clinical trial.*
2. $\frac{\partial \hat{b}(x_i;h^T)}{\partial \bar{x}_T(x_i)} > 0$*: the perceived benefit of a treatment to a patient is increasing in the representation of patients with similar characteristics in the clinical trial.*
3. $\frac{\partial^2 \hat{b}(x_i;h^T)}{\partial \bar{x}_T(x_i)^2} < 0$*: the degree to which increasing representation in a clinical trial positively impacts perceived benefits for group members is decreasing in the group's existing trial representation.*

The intuition is straightforward: when a treatment works better than expected in the trial, people update their beliefs upwards on treatment efficacy.[26] But the degree to which they update depends on the (effective) sample size of the trial. Given that people place positive probability on characteristic $x_i$ mattering, the effective sample for patients with characteristics $x_i$ is increasing in their trial representation. Diminishing returns to representation follows from diminishing returns to sample size in (*e.g.*, Bayesian) models of updating.[27]

We assume posteriors from the most similar previous treatment become the prior for a novel drug.[28,29] That is, letting the most similar past treatment to $T$ come in period $Z < T$, $\alpha(x_i;h^{T-1}) = \alpha(x_i;h^Z)$, $\beta(x_i;h^{T-1}) = \beta(x_i;h^Z)$, $\alpha(h^{T-1}) = \alpha(h^Z)$, and $\beta(h^{T-1}) = \beta(h^Z)$. Given this assumption, even when all groups begin with the same prior beliefs on efficacy at the beginning of time (in period 0), the underrepresentation of a given group will lead to a divergence in the perceived benefit of treatment over time.[30] This divergence has important implications for behavior, described next.

### III.1.2 Patient and Doctor Behavior

Suppose that a patient with characteristics $x_i$ participates in a trial for treatment $T$ when she is invited to participate and

$$\hat{b}(x_i;h^{T-1}) - n_T^{trial} + \varepsilon_{iT}^{trial} \geq 0,$$

where $n_T^{trial}$ equals the non-price costs of participating in the trial (or convincing a patient to do so) and $\varepsilon_{iT}^{trial}$ is a stochastic shock that is i.i.d. across $i$ according to a cumulative distribution function $F_\varepsilon(\cdot)$.

[26] Proofs can be found in the Appendix Section F.3.

[27] For a numerical example, see Appendix F.5.

[28] Similar treatments could, for example, refer to treatments in the same category (drug class), or potentially all treatments for the same disease.

[29] Our analysis would be unchanged qualitatively if people's priors were constructed as a weighted average of their posteriors regarding previous treatments, with more similar treatments receiving larger weights, or if priors were constructed through a simulation mechanism akin to that modeled by Bordalo et al. (2022).

[30] See Appendix F.5 for an example of this divergence.

Similarly, after a successful trial, a patient is treated for treatment $T$ when indicated and

$$\hat{b}(x_i;h^T) - n_T - p_T + \varepsilon_{iT} \geq 0,$$

where $n_T$ refers to the non-price costs of prescribing or adhering to treatment $T$, $p_T$ is the price (*i.e.*, copay) for $T$, and $\varepsilon_{iT}$ is a stochastic shock that is i.i.d. across $i$ according to $F_\varepsilon(\cdot)$. Let

$$d(x_i;h^{T-1}) = \Pr\left(-\varepsilon_{iT}^{trial} \leq \hat{b}(x_i;h^{T-1}) - n_T^{trial}\right)$$

be the likelihood that a patient with characteristic $x_i$ participates in a trial when invited. Similarly, let

$$d(x_i;h^T) = \Pr\left(-\varepsilon_{iT} \leq \hat{b}(x_i;h^T) - n_T - p_T\right)$$

be the likelihood a patient with characteristic $x_i$ is treated for treatment $T$ when the treatment is indicated.

**Corollary 1.** *Given Proposition 1, a patient's demand to participate in a given trial (or a physician's decision to recommend a trial) is increasing in the degree to which patients who shared their (their patients') characteristics were represented in previous trials Z for which the average trial efficacy exceeded prior-belief ratios. Formally, for such trials Z,*

$$\frac{\partial d(x_i;h^{T-1})}{\partial \bar{x}_Z(x_i)} > 0.$$

This result implies that a failure to represent groups in a trial today creates an *intertemporal externality*, as it becomes more difficult to recruit those groups in a trial tomorrow. Such less-represented group members perceive limited benefits from novel treatments relative to members of more-represented groups.

Appendix F.2 formalizes two additional results on how beliefs impact behavior. First, Corollary 3 shows that the comparative statics Proposition 1 establishes for beliefs also hold for behavior: the demand for a new medication is increasing in the efficacy observed in the clinical trial and the representation of patients with similar characteristics in the clinical trial, with diminishing returns to the latter.[31] Second, Corollary 4 shows how historical and contemporaneous underrepresentation of Black patients in clinical trials creates a gap in the perceived benefits and demand for novel drugs between White and Black patients, where White patients have higher perceived benefits and demand relative to Black patients. It goes on to show how increasing Black representation in clinical trials closes these gaps. But this begs a question that requires analyzing the decisions of pharmaceutical firms: Why do these gaps persist?

[31] The last result on diminishing returns requires mild regularity conditions on $F_\varepsilon(\cdot)$.

## III.2 Pharmaceutical Firms

Pharmaceutical firms choose a trial-*r*ecruitment strategy $r \in R$ to maximize profit

$$\Pi_r = s_r \times v_r - c_r,$$

where $s_r \in [0,1]$ is the success probability of the trial, $v_r \geq 0$ is the value of a successful trial to the firm, and $c_r \geq 0$ is the cost to the firm of running the trial. Both the value and cost to the firm of recruitment strategy $r$ depend on patients' and doctors' assessments of the benefits of treatment, which in turn influence trial-participation and treatment decisions.

### III.2.1 The Value and Cost to the Firm of Increasing Trial Representation

The behavior of physicians and patients described in Section III.1 influences firms' recruitment decisions. Suppose firms have access to a status-quo technology for recruiting patients to clinical trials, which firms use to *invite* Black and White patients to participate in the trial in proportions $\bar{x}_T^a$ and $(1-\bar{x}_T^a)$, respectively. The actual trial representation of these groups under the status quo, $\bar{x}_T^{sq}$ and $(1-\bar{x}_T^{sq})$, generally differs from the invited proportions — and this may vary by $x_i$ (*i.e.*, there may be differences in accrual rates across groups.)

**Proposition 2.** *Let* $d(x_i;h^{T-1}) = \Pr\left(-\varepsilon_{iT}^{trial} \leq \hat{b}(x_i;h^{T-1}) - n_T^{trial}\right)$ *be the likelihood a patient with characteristic* $x_i$ *participates in a trial when invited. Then, the share of Black trial participants under the status quo recruitment technology is given by:*

$$\bar{x}_T^{sq} = \frac{d(x_i=1;h^{T-1}) \times \bar{x}_T^a}{d(x_i=1;h^{T-1}) \times \bar{x}_T^a + d(x_i=0;h^{T-1}) \times (1-\bar{x}_T^a)}.$$

**Corollary 2.** *Proposition 2 implies that Black trial representation will be lower than its invited representation under the status quo technology when the demand for trial participation of Black patients falls below that of White patients. Formally,* $\bar{x}_T^{sq} < \bar{x}_T^a$ *when* $d(x_i=1;h^{T-1}) < d(x_i=0;h^{T-1})$.

Under the status quo technology, a racial gap in perceived treatment benefits increases the racial gap in trial participation. Firms could choose to incur a cost to increase Black representation from its level under the status quo, and there would be value to them doing so: due to diminishing returns to representation (Proposition 1), it could increase demand among Black patients and their doctors without much harm to White patients or their doctors. That is, there's value to firms investing in *inclusive infrastructure*, a term we develop and provide concrete examples of in Section VI. However, the trade-off involved with increasing participation of historically underrepresented groups is that investing in such infrastructure is costly. Specifically, we assume it requires a fixed cost $f > 0$ and, due to the mechanism highlighted in Corollary 2, the returns to such investment are not completely internalized by any given firm: It increases perceived benefits for all similar treatments in the future, including those developed by

*other* firms.[32] Thus, firms underinvest in such technology relative to what is socially optimal.[33] The externality a firm's *current* recruitment decisions have on other firms' *future* recruitment costs enables a cycle of underrepresentation.

## III.3 The Cycle of Underrepresentation

**Proposition 3.** *Suppose the most similar treatment $Z$ to $T$ outperformed patients' prior expectations. When the fixed costs $f$ to deviating from the status-quo recruitment technology to inclusive infrastructure are sufficiently large, then underrepresentation of Black patients in the historical trial leads to further underrepresentation of Black patients in the current trial:*

$$\frac{\partial \bar{x}_T}{\partial \bar{x}_Z} > 0.$$

This result flows from the externality described above and illustrated with a numerical example in Model Appendix Section F.5.

Together, *Propositions 1–3* describe a cycle of underrepresentation: (1) Trials in the past have not been representative of Black patients. (2) The lack of representation decreases the perceived benefits of treatments for Black patients and physicians who treat them. (3) The aforementioned (*i.e.*, 1 and 2) make it more costly for firms to actively increase trial representation. (4) Trials today are not representative for Black patients.[34] (5) And the cycle continues. Table II summarizes our theoretical predictions and how they connect to our empirical results, which we turn to next.

[32] Firms may also be able to free-ride on investments made in inclusive infrastructure by the public sector or other firms (reducing fixed-costs $f$), which is an additional channel by which firms wouldn't fully internalize the social benefits of such investments.

[33] See Model Appendix Section F.4 for details.

[34] While Proposition 3 suggests that Black representation could get worse over time in a cycle of underrepresentation, it abstracts from policy efforts to improve representation (see Appendix G.1). We view the proposition as identifying a force that pushes against such policy efforts.

Table II: Summary of Theoretical Predictions and Empirical Results

| Theory | Predictions | Exhibits | Result Summary |
|---|---|---|---|
| Prop. 1.1; Cor. 3.1 | Perceived benefits and demand for a new medication are increasing in trial-reported efficacy. | ***Table III*** | A 1 sd increase in efficacy increases physician prescribing intention by 0.28 sd. |
| Prop. 1.2; Cor. 3.2 | Perceived benefits and demand for a new medication are increasing in representation of similar patients in clinical trials. | ***Table III*** | • For physicians, a 1 sd increase in representation increases prescribing intention by 0.11 sd.<br>• For Black patients, being assigned to the representative treatment increases self-reported relevance for their own care ("relevance") and the likelihood that their posterior on efficacy is within a small neighborhood of the reported clinical-trial results ("loading on the signal") 0.78 sd and 19.9 pp, respectively. |
| Prop. 1.3; Cor. 3.3 | Diminishing returns to representation. | ***Figure II(d); Table III*** | • For Physicians treating White Patients ("PWP"), we fail to reject the null hypothesis that a decrease in White representation (from existing high levels) does not change prescribing intention.<br>• For White patients, we fail to reject the null hypothesis that a decrease in White representation (from existing high levels) does not change relevance or loading on the signal. |
| Cor. 4 | • There are White-Black gaps in perceived benefits and demand for a new medication.<br>• Increasing Black representation in clinical trials narrows these gaps. | ***Figure III; Figure IV; Figure V*** | • PWP have a mean prescribing intention of 6.46 while PBP who are exposed to non-representative trials have a mean prescribing intention of 4.90. The prescribing intention of PBP who are exposed to representative trials increases to 6.26 and is statistically indistinguishable from that of PWP.<br>• Black patients who are shown the low representation trial are 26 pp less likely to load on the signal than White patients. Black patients shown the representative trials are only 5 pp less likely to load on the signal than White patients and this difference is statistically indistinguishable from 0. |
| Prop. 2; Cor. 2 | Groups that were historically underrepresented have a lower propensity to participate in trials today than historically well-represented groups. | **Table I** | Black respondents are 9.9 pp less likely to perceive science as beneficial, 7.9 pp less likely to have heard of clinical trials and 5.4 pp less willing to participate in clinical trials when recommended by a doctor. |
| Prop. 3 | In the absence of government regulation or other public-policy intervention, low representation of Black patients in clinical trials is a persistent equilibrium outcome. | **Figure I(a)**; SectionVI.1 | • Black participation in pivotal trials remains low at a median of five percent over time. In the absence of investment in inclusive infrastructure (relative to in HIV/AIDS trials), cancer trials had five percent Black representation.<br>• By contrast, public involvement and investment in inclusive infrastructure for HIV/AIDS trials was associated with 28 percent Black representation. |

*Notes*: Formatting of the exhibits indicate the type of evidence: ***causal evidence***; **descriptive evidence**; suggestive evidence.

# IV Experimental Design

## IV.1 Experimental Design

To test several of these predictions, we conducted two survey experiments — one with a sample of primary care physicians, and one with a sample of patients.[35] The experiments differed in important ways reflective of the different subject pools. Physicians, who are familiar with evaluating new medications as part of their practice, were asked to rate several hypothetical drugs. In each drug profile, we cross-randomized the racial composition of the trial and efficacy.[36] Drug efficacy was used as a "numeraire" since it is widely considered the most important characteristic of a new medication. Prescribing intention and relevance for own patients for each medication were assessed. When surveying patients, we presented a simpler exercise: respondents were shown trial evidence associated with a single actual drug. Primary outcomes for patients included beliefs on the drug's efficacy, relevance for own health, and willingness to "ask their doctor" about the new medication.[37] We describe the experiments immediately below and discuss common critiques of survey experiments as well as how we endeavored to overcome them in Subsections V.2.4 and V.2.5, respectively.

### IV.1.1 Physician Survey Experiment

We recruited physicians who met the following criteria: (i) actively practicing in primary care, (ii) practicing in an outpatient setting,[38] and (iii) holding either an MD or DO. We worked with a licensed vendor of the American Medical Association's (AMA) physician masterfile to identify and contact eligible physicians. We verified that survey respondents met all three criteria with a set of screening questions at the outset of the experiment. We pre-specified that the representativeness of the trial sample could interact positively with the demographic composition of the physician's patient panel. Thus, to ensure suitable variation in the panel, we split zip codes into deciles by Black population, weighting each zip code by its total population, and requested that half of all physician contacts be pulled from the top decile, one-quarter from the bottom decile (these two deciles account for 15 percent of all primary care physicians), and one-quarter from the remaining deciles.[39] This sampling approach was motivated by the fact that the distribution of Black patients across geographies and providers tends to be highly concentrated (Bach et al. 2004; Chandra, Frakes and Malani 2017).

[35] Appendix Figure B9 depicts the flow of the physician and patient surveys.

[36] We used hypothetical drugs instead of real drugs since there were not nearly enough real-world trials to experimentally include a range of Black patients and carefully titrated mechanisms of action and efficacy. Such an approach of using hypothetical drugs was followed by Kesselheim et al. (2012) to measure the influence of the source of clinical trial funding on the prescribing behavior of doctors.

[37] This language was chosen intentionally to mirror standard DTCA in the U.S., one of the primary contexts in which patients engage, unassisted by a physician, with medical information.

[38] We excluded hospitalists from our sample.

[39] We determine zip code rank using 5-year zip code-level population estimates reported in the 2019 American Community Survey.

We sent each physician a personalized email (to their professional email address) inviting them to participate in a study. The email originated from a Harvard email account. We embedded a message as email text, which noted that the purpose of the study was to collect physician views on clinical trials research, that the study had received IRB approval, that their data would be securely stored, and that the study was not funded by industry but rather for academic purposes (see Appendix Exhibit E1). The letter explained that the physician respondents would be asked to rate eight hypothetical drugs and would be compensated $100 for their participation.[40]

Although the vignettes were hypothetical, the drugs were based on recently developed therapies to treat diabetes. We chose to focus on diabetes because it is a common condition that is typically managed by primary care providers, and several new therapies with novel mechanisms of action have recently been developed (American Diabetes Association 2020). Additionally, there are no established guidelines or biomedical findings that would justify differential treatment on the basis of race or ethnicity (Golden et al. 2012).

After confirming eligibility and answering questions about their practice, physicians were shown eight unique drug profiles. Profiles were selected randomly without replacement (*i.e.*, physicians never saw an exact duplicate) and drug names were selected from 15 alternatives.[41] At the top of each profile, we listed the generic name of a hypothetical drug, which we developed by following standard naming conventions (*e.g.*, suffixes and prefixes) that convey information about a drug's type. Profiles also included the drug's mechanism of action, the study type, sample size, and sample demographics (see Appendix Exhibit E2 for an example of a profile and Appendix Exhibit E3 for a table listing the hypothetical drugs shown to participants). Profiles were randomly assigned an efficacy value ranging uniformly from a 0.5–2.0 percent average reduction in A1c, conforming to typical values of FDA-approved oral antiglycemics (Wexler 2022; Nathan et al. 2009), and a percent Black of trial subjects value ranging from 0–35 percent, with lower values oversampled as trial diversity is typically low (Knepper and McLeod 2018; Dornsife et al. 2019).[42] Note that only efficacy and percent Black varied across the profile, with all else held fixed.[43] In each case, the trial type was listed as a double-blind active comparator trial and the sample size was fixed at 1,500 participants.[44]

[40]We piloted this survey with $75 honoraria but raised compensation to increase yield. The only meaningful deviation from our pre-analysis plan was that we planned to recruit 1000 hypertensive patients, but it proved difficult to find that many who met both our demographic and medical criteria in the online survey panel (Banerjee et al. 2020).

[41]There were 8,640 unique profiles: 15 hypothetical drugs multiplied by 16 possible efficacy values (0.5–2.0 percent reductions in A1c in 0.1 percent increments) multiplied by 36 possible values of percent Black of trial subjects (0–35 percent in 1 percent increments).

[42]Values of percent Black ranging from 0-4 percent were sampled with probability 0.33, values ranging from 5–14 percent were sampled with probability 0.34, and values ranging from 15–35 percent were sampled with probability 0.33.

[43]Appendix Table C2 demonstrates that both the mean and the range of representation and efficacy values assigned to physicians are uncorrelated with a host of physician individual and patient panel characteristics.

[44]We did not also provide statistics on breakdown by gender in the drug profile. Although sex is another important feature of identity, introducing this information would require respondents to weigh a variable that unequivocally affects drug metabolism (see, for example, the NIH's "sex as a biological variable" https://www.nih.gov/news-events/videos/considering-sex-biological-variable-clinical-trials). Moreover, the policy issue of underrepresentation of women in trials is not as acute (see Appendix Figure B3).

After viewing each profile, physicians were asked to rate how relevant the findings from the trial were for their patients (akin to the EBM step) and how likely they would be to prescribe the drug for patients with poorly controlled diabetes in their care. Both outcomes were on a scale from 0 to 10.[45] After reviewing all drug profiles, respondents were asked about their confidence in extrapolating trial findings across demographic groups or geographies. In the final survey section, we asked questions about risk aversion, time preference, and altruism and posed open-text questions used in sentiment analyses.

We sent a follow-up survey to physicians one to three weeks after they initially completed the survey. In the follow-up survey, we allocated $5 to each physician and asked how they would like to divide the amount between two real-world campaigns supporting recruitment efforts for clinical trials (see Appendix Exhibit E6). The first campaign aimed to boost trial participation among the American public at large, while the second campaign aimed to boost trial participation for underrepresented minority communities.[46]

### IV.1.2 Patient Survey Experiment

Patients were recruited from Lucid, an online survey platform frequently used in social science research and marketing. Respondents were told that the survey was designed to solicit their views on health care and to understand the factors that affect their interest in health research. Eligibility criteria for passing the screening questions included: (1) self-reported non-Hispanic White or non-Hispanic Black race/ethnicity, (2) at least age 35, and (3) endorsement of a diagnosis of high blood pressure (alone or comorbid with other conditions). To verify that respondents had, in fact, been diagnosed previously with hypertension, they were asked to enter their latest systolic and diastolic blood pressure readings in an open-text field.[47] Any respondent entering nonsensical values for blood pressure was deemed ineligible. We focused on high blood pressure instead of diabetes because a larger share of adults in the U.S. suffer from hypertension (45 percent) than diabetes (15 percent), thus facilitating recruitment (Ostchega et al. 2020; Center for Disease Control and Prevention 2021). We introduced consequentiality by explicitly encouraging patients to answer truthfully, and noting their responses would be used to generate a personalized report they could download and share with their health provider. Approximately 42 percent of patient respondents downloaded the personalized report.

We did not assume patients were familiar with clinical trials, so began the experimental module by providing basic details about the process. Before randomization, we told respondents that new medications to treat blood pressure are studied by researchers all the time.[48] We noted these new therapies typically aim to improve blood pressure control, reduce complexity, or decrease side effects

[45] The exact question wording shown to physicians can be found in Appendix Exhibit E5.

[46] Both campaigns were run by a non-profit, the Center for Information and Study on Clinical Research Participation (CISCRP).

[47] By declining to provide a range of values or a dropdown menu, we screened out any individuals who were unfamiliar with the scales for either measurement and thus less likely to carry the diagnosis.

[48] A link to the questionnaire is found in Appendix Exhibit E7.

from medication. We added that new medications may not be an improvement over previous therapies, and thus must be tested before they are widely available. Patients were then shown details about a new medication: a combination antihypertensive medication. We asked each patient whether they had heard of the new drug before (95 percent had not) and what they anticipated the effect of the medication would be on their systolic blood pressure (in units of millimeters of mercury [mmHg]).

Patients were then provided with findings from an actual clinical trial. We randomly assigned respondents to see trial data from studies that enrolled different shares of Black patients. The medication we presented was tested in two separate locations: in one setting, the percent Black in the trial was less than one percent — approximately one-third of trials in the ClinicalTrials.gov database meet such a criterium — and in the second, the percent Black in the trial was 15 percent. Efficacy was strong and comparable in both settings, lowering systolic blood pressure by about 15 mmHg.[49] We thus randomized only the percent Black in the trial, holding efficacy and all other parameters of the trial constant.

After being shown information on the drug's efficacy and the randomized racial composition of the study, patient respondents were again asked to provide their beliefs about the drug's efficacy. Additionally, respondents reported how relevant the findings of the trial were to patients like them and whether they would be interested in "asking their doctor" about the medication.[50] We also asked patients the same question we had posed to doctors about extrapolating from trials generically. If patients indicated that they were not confident in extrapolating, we asked them to indicate their rationale.

In the final sections of the survey, we inquired about trust, risk aversion, altruism, and time preferences. We also asked respondents to describe their current primary health care provider and current regimen for blood pressure management medication adherence. We concluded with open-text questions and a link to learn more about clinical trials.

# V Experimental Data and Results

## V.1 Sample Characteristics

We invited 12,192 physicians to participate in the study.[51] Amongst those who passed the screening questions, 87 percent completed the survey (137 physicians); completion rates did not vary significantly across strata. Potential respondents were most commonly screened out if they were not practicing primary care physicians, or if they were hospitalists (*i.e.*, not outpatient providers). On nearly all

[49]To hold efficacy precisely constant across trials, we reported to participants that treated subjects in their assigned trial saw their systolic blood pressure drop significantly compared to subjects in the control group, and then stated that across similar studies the average drop in systolic blood pressure among participants taking the medication was about 15 mmHg.

[50]The exact question wording shown to patients can be found in Appendix Exhibit E8.

[51]In total, 4.7 percent of emails bounced and 1.8 percent of those invited started the survey, in line with standard click-through rates in the marketing literature (Richardson, Dominowska and Ragno 2007; Kanich et al. 2009). Our click-through rate of 1.8 percent was considerably higher than the 0.25–0.5 percent quoted to us by vendors as typical for email marketing campaigns.

dimensions, the characteristics of physicians in our sample are comparable to those of physicians in the same zip code strata in the AMA Masterfile (see Appendix Table C3), with the following exceptions: sample physicians from the top Black share decile stratum tend to be older and from higher ranked medical schools, and physicians in other zip codes tend to have a higher share White population and a lower share Hispanic population.[52]

We recruited 275 patients diagnosed with hypertension: 139 Black and 136 White respondents. Respondents are comparable to individuals with hypertension in the Medical Expenditure Panel Survey; in Appendix Table C5, we document that Black respondents in our survey are broadly similar to Black MEPS respondents by age, sex, geography, income, and insurance status, although they have slightly higher levels of college education (Blewett et al. 2019). White survey respondents exhibited similar sex, geography, and insurance characteristics to White MEPS respondents, though were slightly older and less educated than those in the MEPS data.[53] There was no significant imbalance or differential attrition across arms for either survey (see Appendix Tables C2, C7, C8, and C9).

## V.2 Estimation and Results

To test whether increasing representation of Black patients (which we refer to simply as "representation") in trials affects how physicians view study results and make prescribing decisions, we estimate the following equation:

$$Y_{jk} = \alpha_0 + \alpha_1 \text{Representation}_{jk} + \alpha_2 \text{Efficacy}_{jk} + \rho_k + \mu_j + \sigma_{jk} + \varepsilon_{jk}, \tag{2}$$

where $j$ denotes a drug and $k$ denotes a unique physician respondent, $Y_{jk}$ denotes our primary outcomes of interest, relevance for one's own patients, and willingness to prescribe. Representation is the share of patients in a given trial who are Black. Efficacy captures the percentage point drop in measured hemoglobin A1c. Both efficacy and representation were cross-randomized in each profile. Our pre-specified main estimating equation includes physician fixed effects ($\rho_k$), drug mechanism fixed effects ($\mu_j$), and indicators for the order in which profiles were shown ($\sigma_{jk}$), though we also present results without any controls. The outcome and randomized attributes are standardized. Standard errors are clustered at the physician level. We also pre-specified heterogeneity, interacting trial demographics with those of the doctor's panel, hypothesizing that if representation mattered to doctors, it would matter more for those that treat Black patients.

To test whether the racial composition of clinical trials affects patient beliefs and behavior, we estimate for patient $i$ of race $r$ the following:

[52] Approximately 60 percent of the physicians who completed the initial survey responded to the follow-up email. The physicians who responded to the follow-up survey were comparable to those who did not respond to the follow-up survey (see Appendix Table C4).

[53] Our main results are robust to including person weights derived from a nationally representative survey, the Medical Expenditure Panel Survey (Appendix Table C6).

$$Y_{i(r)} = \beta_0 + \beta_1 1_{i(r)}^{\text{Representative}} + X'_{i(r)}\Omega + \varepsilon_{i(r)}, \tag{3}$$

where the indicator variable captures the difference between receiving the information that the percent Black of trial participants was 15 percent versus less than 1 percent. Recall that efficacy was held fixed, and all respondents saw the same drug. We estimate Equation 3 separately by patient race for three outcomes: relevance, efficacy beliefs, and asking one's doctor. Relevance (of the drug for me) is transformed from a Likert scale (0 to 10) to standard deviation units. Loading on Signal is an indicator equal to one if patients' beliefs about personal efficacy are within 1 mmHg of the reported treatment effect in the trial.[54] Ask Doctor is an indicator variable equal to one if patients indicate a desire to talk to their doctor about the drug.

### V.2.1 Main Findings

Table III presents our main results for both experiments: Panel (a) reports findings for physicians and Panel (b) for patients. Panel (a) Columns (1) and (2) include only the randomized components of drug profiles. A one standard deviation increase in the reported efficacy of the drug — a reduction in A1c of roughly 0.44 percentage points — increases relevance and willingness to prescribe a medication by 0.165 and 0.229 standard deviation units, respectively. Conditional on the drug's efficacy, a one standard deviation increase in percent Black — about a 10 percentage point increase in Black trial participants — increases relevance for patients by 0.163 standard deviation units and willingness to prescribe the drug by 0.179 standard deviation units. Columns (3) and (4) present our main specification (Equation 2). We find representation affects both relevance and intent to prescribe, increasing both by approximately 0.11 standard deviation units.

The *p*-values displayed in the bottom rows of these last two columns indicate that — although we reject that the coefficients on representation and efficacy are equal — we *cannot* reject that representation has about *half the effect* of efficacy. In other words, physicians are approximately half as responsive to who was in the trial as they are to how well the drug works. The results in Columns (5) and (6) — in which we include interaction terms between experimentally-manipulated measures of representation and efficacy with each physician's Black patient share — are key in understanding our results: the effect of increased Black representation on prescribing behavior is attributable to doctors that treat at least *some* Black patients. We observe no comparable (significant) interaction between doctors' patient demographics and efficacy.

In Table III, characteristics of the physician's patient panel are entered linearly. Figure II explores these

[54] Non-standardized outcomes and continuous updating outcomes yield similar results, which are gathered in the Appendix Table C10. Note that our approach deviated from many tests of Bayesian updating in that we did not vary the signal on drug efficacy (Hjort et al. 2021; Jensen 2010; Roth and Wohlfart 2020). Rather, the intervention informed patient respondents of a distinct feature of the data-generating process — the composition of the sample — that our framework predicts influences the weight they place on the signal in assessing how much the drug would personally benefit them. Our focus is then on this weight, as measured by whether patients' posterior beliefs were within 1 mmHg of the reported signal.

relationships nonparametrically by interacting quantiles of patient percent Black with the treatment and plotting the total effect (main effect + interaction). Panel (a) shows the results for efficacy, demonstrating a relatively constant effect on relevance and prescribing across the percentage Black of patients. By contrast, in Panel (b) representation has a nearly linear and upward-sloping relationship: the higher percentage Black in a doctors' patient panel, the more they respond to the inclusion of Black patients in the trial.[55] Note that this line naturally begins at zero over the domain we test: there is simply a null effect (not a strong negative effect) of increasing Black representation among physicians who care mostly for White patients.

To provide further assurance that it is indeed specifically the racial composition of the panel that is driving the heterogeneity, Appendix Figure B11 presents an omnibus test, in which physician-specific representation coefficients are regressed on panel demographic characteristics. A significant association exists only between the magnitude of the coefficient and the panel percent Black, with no strong relationship between representation and percent female, Hispanic, foreign-born, or senior citizen. Moreover, there is no significant relationship between physician-specific efficacy coefficients and percent Black, nor between the other demographic categories.

We next turn to findings from patients in Panel (b) of Table III. Recall that in this specification (Equation 3), the treatment is an indicator variable. We split the sample by patient race, with findings from Black patients displayed in the odd columns and results from White patients shown in the even columns and a $p$-value of the difference between the two samples in the bottom even rows. Column (1) reports that Black patients with hypertension assess clinical trials with 15 percent Black participants as 0.781 standard deviation units more relevant than trials with less than 1 percent Black participants — holding drug name, mechanism, and reported efficacy constant. This result is statistically significant at the 1 percent level. Column (3) indicates that these higher assessments translate into a positive but statistically insignificant willingness to ask their physician about the medication. Column (5) reports that the representative arm is associated with a 19.9 percentage point increase in believing the drug would perform as well on oneself as in the trial. The results from White patients with hypertension are mixed in sign and never statistically significant (Columns 2, 4, and 6).

Results from our patient sample are also broadly consistent with the model's prediction of diminishing returns to representation: representation matters for Black hypertensive patients, and does not (over the domain tested) for White patients, similar to what we find for prescribing intentions in Figure II. Taken together, the results suggest physicians are acting as good agents for their patients — combining the evidence on efficacy while also taking patient views into account (Ellis and McGuire 1986; Barnato 2017).

[55] Similar results are obtained when using deciles of percent Black (Appendix Figure B10).

Table III: Physician and Patient Experimental Results on Effects of Increasing Representation

**Panel A: Primary Care Physicians**

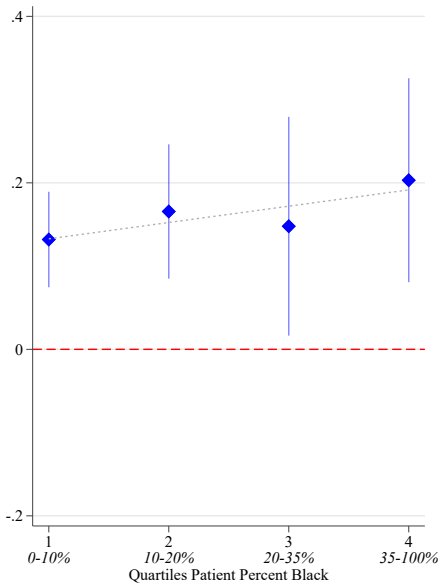
| | *Relevance* | *Prescribing* | *Relevance* | *Prescribing* | *Relevance* | *Prescribing* |
|---|---|---|---|---|---|---|
| | No Controls | | Main Specification | | Share Black Interactions | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Representation (Own Group) | 0.163*** | 0.179*** | 0.109*** | 0.107*** | 0.007 | -0.005 |
| | (0.039) | (0.036) | (0.029) | (0.029) | (0.038) | (0.039) |
| Efficacy | 0.165*** | 0.229*** | 0.189*** | 0.281*** | 0.179*** | 0.285*** |
| | (0.038) | (0.039) | (0.029) | (0.032) | (0.036) | (0.043) |
| Representation × Patient Percent Black | | | | | 0.004*** | 0.004*** |
| | | | | | (0.001) | (0.001) |
| Efficacy × Patient Percent Black | | | | | 0.001 | -0.001 |
| | | | | | (0.001) | (0.001) |
| *p*-value: Representation=Efficacy | | | 0.057* | <0.001*** | | |
| *p*-value: Representation=$\frac{1}{2}$(Efficacy) | | | 0.655 | 0.314 | | |
| Doctor FEs | No | No | Yes | Yes | Yes | Yes |
| Profile Order FEs | No | No | Yes | Yes | Yes | Yes |
| Rx Mechanism FEs | No | No | Yes | Yes | Yes | Yes |
| Observations | 1,096 | 1,096 | 1,096 | 1,096 | 1,096 | 1,096 |

**Panel B: Patients**

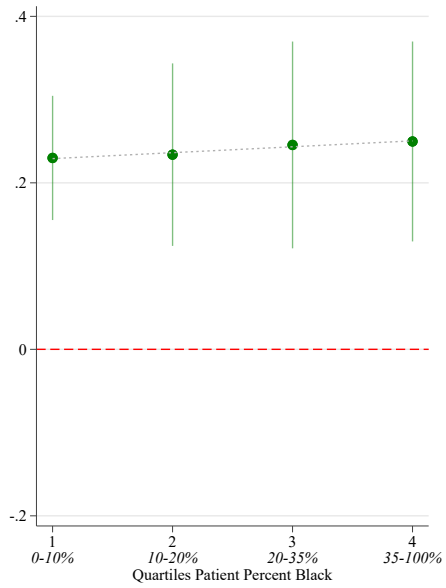
| | *Relevance* | | *Ask Doctor* | | *Loading on Signal* | |
|---|---|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Representative Treatment | 0.781*** | 0.172 | 0.021 | 0.006 | 0.199** | -0.057 |
| | (0.167) | (0.159) | (0.077) | (0.079) | (0.083) | (0.086) |
| Control Mean | -0.26 | -0.23 | 0.70 | 0.70 | 0.33 | 0.59 |
| Observations | 139 | 136 | 139 | 136 | 139 | 136 |

*Notes:* Panel (a) reports OLS estimates for the outcomes of *Relevance* and *Prescribing Intention* on the sample of primary care physician respondents. *Representation* refers to the randomized percent Black in the trial unless otherwise indicated. *Efficacy* refers to the randomized percentage point drop in A1c. *Prescribing Intention*, *Representation* and *Efficacy* are standardized to a mean of 0 and a standard deviation of 1. Columns (3) and (4) report results from the main specification (Equation 2). Columns (5) and (6) interact *Representation* with the reported percent of patients that are Black in the physician's panel, the main effect is included but not reported. 137 physicians participated in the experiment each assessing eight oral antiglycemic medications. Standard errors clustered at the physician level are in parentheses. Panel (b) reports OLS estimates from Equation 3 on the sample of patient respondents. *Relevance* refers to relevance for own care and is standardized to a mean of zero and standard deviation of one. *Loading on Signal* is an indicator equal to one if the respondent's posterior was within 1 mmHg of the signal (*i.e.*, between 14 and 16) and zero otherwise. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

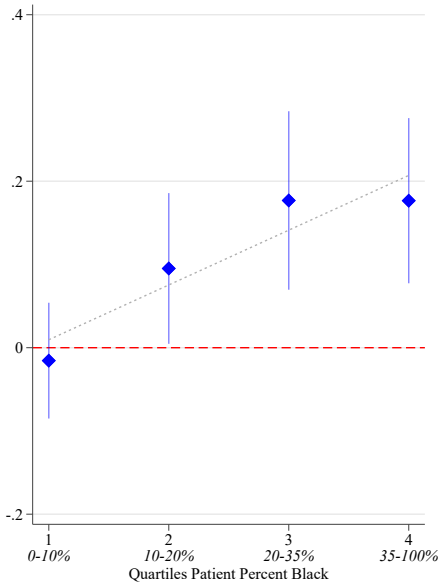
Figure II: Heterogeneity Among Physicians by Racial Composition of Patient Panel



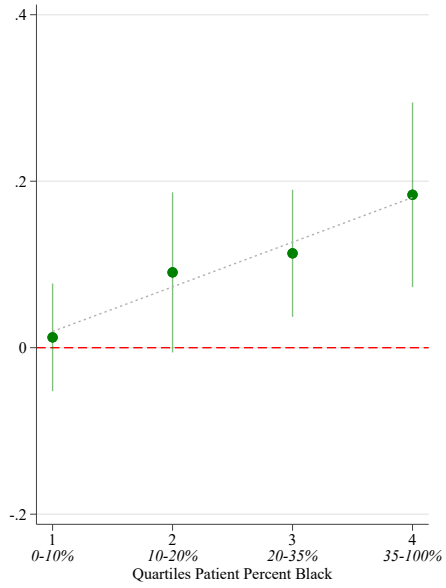

(a) Efficacy on Relevance




(b) Efficacy on Prescribing Intention




(c) Representation on Relevance




(d) Representation on Prescribing Intention

*Notes:* Figure plots OLS estimates for two outcomes — *Relevance* (Panels a and c) and *Prescribing Intentention* (Panels b and d) — from specifications estimated with interaction terms between each quartile of patient percent Black and either *Representation* or *Efficacy*. Fixed effects are residualized before estimating Equation 2. Figure plots the linear combination of the main effect and the interaction with each quartile; quartile one is defined as the reference. Robust standard errors are clustered at the physician level. 95 percent confidence intervals are displayed.

Lastly, we turn to heterogeneity in the patient results in Table IV. Given prior research on the legacy of discrimination and the importance of trust (Alsan and Wanamaker 2018; Alsan, Garrick and Graziani 2019; Greenwood et al. 2020; Gruber and Frakes 2022), we investigated whether the effects of our

intervention would interact with expectations of others' trustworthiness. We measured this using the General Social Survey's question on general trust: "Generally speaking, you would say that: 'Most people can be trusted' or 'Most people cannot be trusted' " (Glaeser et al. 2000; Fehr 2009; Sapienza, Toldra and Zingales 2013; Smith et al. 2021) and saturated our treatment variable. Overall, fewer Black respondents (45%) than White respondents (60%) answer that most people can be trusted. Column (3) demonstrates that the interaction between being treated and expected trustworthiness has a sizeable effect on the willingness to ask a doctor about the drug. Put differently, even if Black patients believe findings are relevant *and* update in the expected direction on efficacy (Columns 1 and 5), there may still be a barrier from inquiring about the medication from their primary care physician if they do not anticipate that others merit their trust.

Table IV: Heterogeneity Among Patients by Expectation of Trustworthiness

| | *Relevance* | | *Ask Doctor* | | *Load on Signal* | |
|---|---|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Treatment x (Expt. Trust.=1) | 1.049*** | 0.190 | 0.291*** | -0.000 | 0.190 | -0.171 |
| | (0.236) | (0.209) | (0.104) | (0.099) | (0.123) | (0.109) |
| Treatment x (Expt. Trust.=0) | 0.562** | 0.141 | -0.211* | 0.011 | 0.211* | 0.115 |
| | (0.235) | (0.249) | (0.108) | (0.132) | (0.113) | (0.136) |
| Expt. Trust. | -0.276 | 0.060 | -0.142 | 0.089 | -0.032 | 0.159 |
| | (0.269) | (0.245) | (0.115) | (0.116) | (0.117) | (0.122) |
| $p$-value: Expt. Trust. 1 = 0 | 0.146 | 0.880 | 0.001 | 0.947 | 0.901 | 0.104 |
| Observations | 139 | 136 | 139 | 136 | 139 | 136 |

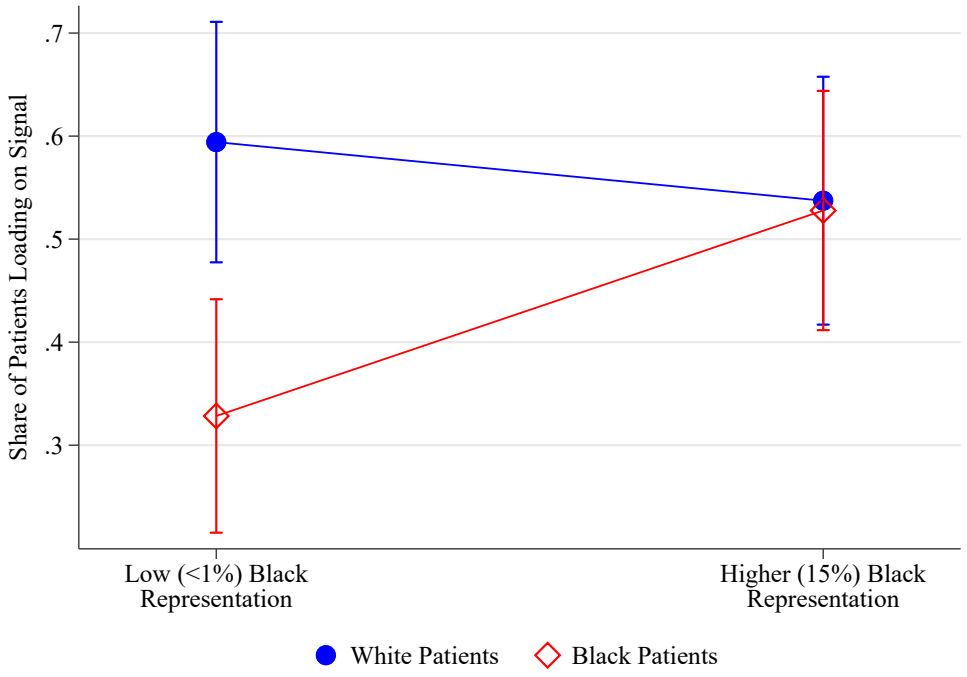
*Notes:* Table reports the OLS estimates for the outcome of *Relevance*, *Ask Doctor*, and *Load on Signal* on the interaction with trust and the main effect for Black and White patients. *Expectation of Trustworthiness* represents the patients' response to the question, "Generally speaking, you would say that: 'Most people can be trusted' or 'Most people cannot be trusted."' *Relevance* is standardized to a mean of 0 and standard deviation of 1. *Load on Signal* and *Ask Doctor* are binary. Columns (1), (3), and (5) report the estimates for Black patients, and Columns (2), (4), and (6) report the estimates for White patients. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

### V.2.2 Representation and Inequality

We next assess whether increased racial representation in clinical trials can close gaps similar to those documented in Figure I. Figure III documents that — when the share Black of the trial is low — a gap emerges between Black and White patients shown identical information on drug efficacy. Black hypertensive patients' views on how much the drug will lower their blood pressure are within 1mm of the range of the reported clinical effect 33 percent of the time, compared to almost 60 percent of the time for White hypertensive patients. This difference is large and statistically significant. When the trial is more

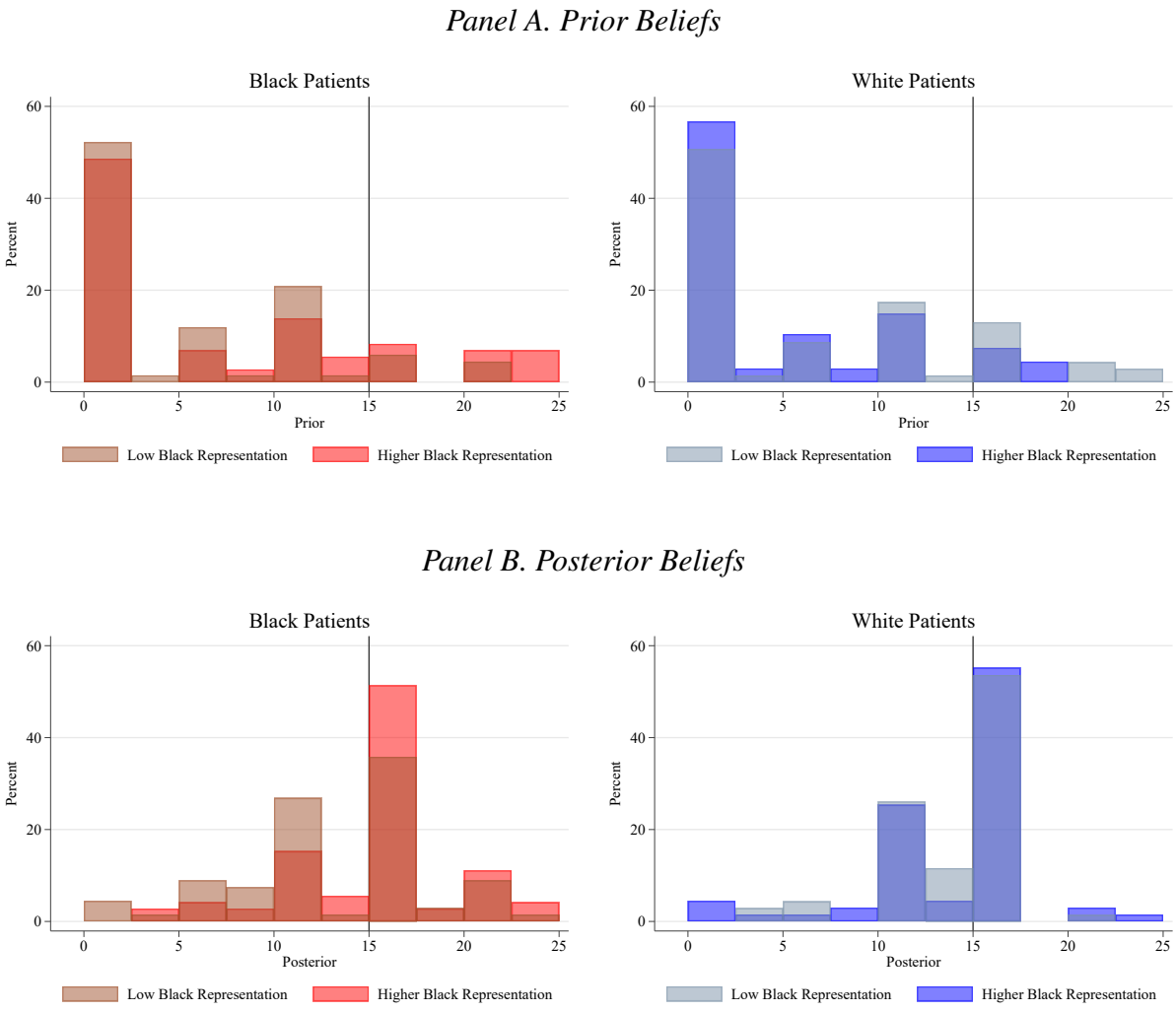
inclusive of Black patients, however, this gap closes. While the change for Black patients is dramatic, the effect on White patients is negligible. This result is also observed when plotting the distributions of prior and posterior views on drug efficacy — the latter under the different interventions. Before the information treatment, the prior distributions for Black and White patients are indistinguishable (see Figure IV, K-S test $p$-value = 0.960). Regardless of the trial arm they are assigned, White patients update vigorously on trial results, reporting an efficacy for their own health that is similar to the study finding. In contrast, Black patients map efficacy into effectiveness for own health much more readily when the sample is more representative (K-S test $p$-value 0.026).

Figure III: Loading on Signal by Race and Treatment Status




*Notes:* Figure plots the share of respondents who "Load on Signal" — whose posteriors are within 1 mmHg of the report drug efficacy in our intervention (15 mmHg) — by race and treatment group. *Load on Signal* is an indicator variable that takes a value of one if the respondent's posterior was between 14 and 16, and zero otherwise. The $x-$axis reports values for two groups of respondents: control (low Black representation) and treated (higher Black representation). Results are plotted separately by respondent race. 95 percent confidence intervals are included.

Figure IV: Prior and Posterior Beliefs by Patient Race and Trial Representation



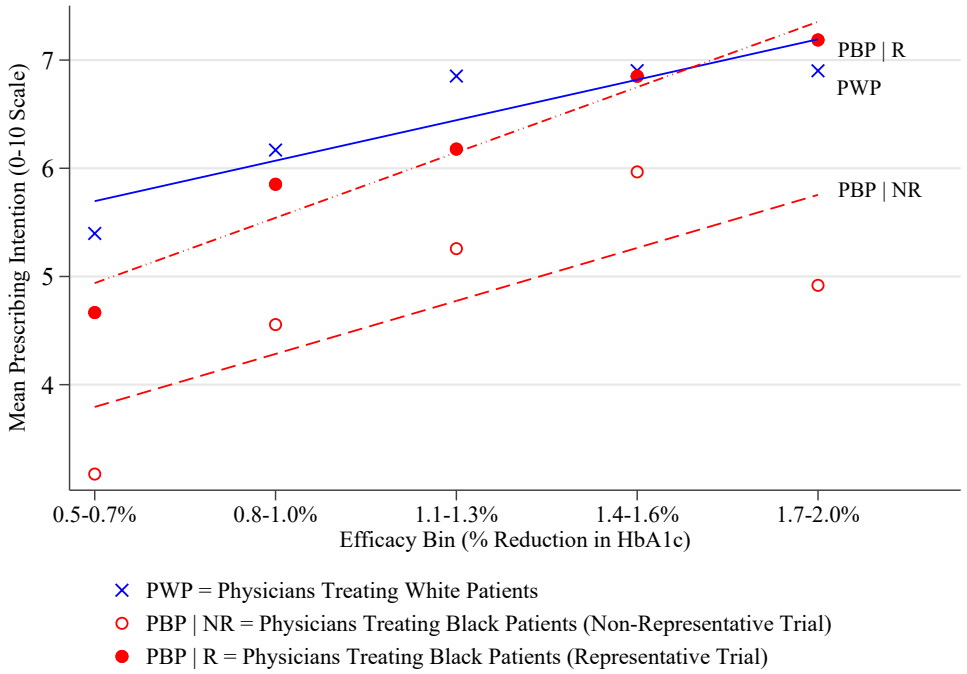

*Notes:* Figure plots the prior and posterior distribution of beliefs about the perceived efficacy of the new antihypertensive medication for the patient's own condition by respondent's race and assigned treatment status (trial shown had either low Black representation or higher Black representation). The signal on efficacy shown to patients (15 mmHg) is displayed as a black vertical line and was revealed to patients following elicitation of priors. A Kolmogorov-Smirnov test fails to reject the null that the priors are identical across race ($p$-value=0.960). For Black patients, Kolmogorov-Smirnov test rejects the null that the posteriors are identical across arms ($p$-value $= 0.026$). For White patients, Kolmogorov-Smirnov test fails to reject the null that the posteriors are identical across arms ($p$-value=0.789).

Our results can be visualized by examining the gaps in prescribing intention across physicians who treat different categories of patients. We divide the sample of physicians into two groups: physicians who treat Black patients (PBP) and physicians who treat White patients (PWP). We define these categories by using the reported characteristics of each physician's reported panel and whether they treat above or below the sample median for the relevant racial group.

Figure V plots prescribing intention across the physician types. Efficacy, as measured by A1c reduction,

is shown on the x-axis and the mean prescribing intention for each efficacy bin is plotted on the y-axis. The upward-sloping line indicates that physicians serving all types of patients are more likely to prescribe medications that were randomly assigned higher rates of efficacy. If a trial has less than 5 percent Black representation (the current median share of Black participation in clinical trials) prescribing intention of physicians treating more Black patients lies below that of physicians treating White patients at every efficacy level. However, when trials become more representative, this gap is erased.

Figure V: Physician Prescribing Intention by Patient Composition and Trial Representation




*Notes:* Figure plots the relationship between *Efficacy* and *Prescribing Intention* (on a 0-10 scale) by patient composition and percent Black of trial subjects in the profiles shown to physicians. *PBP* (Physicians treating Black Patients) denotes physicians who report above the median percent Black patients in their patient panel. /textitPWP (Physicians treating White patients) is defined similarly with respect to White patients. NR indicates non-representative (<5 percent Black in trials) whereas R indicates representative (>= 5 percent Black in trial). Note that 5 percent is the median percent Black in clinical trials (see Figure I).

### V.2.3 Understanding Mechanisms: Extrapolation

Why does representation matter? The model in Section III.1.1 captures the idea that extrapolation from trial data is facilitated by the similarity between patient characteristics and the trial sample. We probe that assumption by asking physicians and patients how confident they are that a drug found to be safe and effective in a study of White patients would be safe and effective for Black patients. Confidence is measured on a scale of 0 to 3 ranging from "Not confident at all" to "High confidence." As such a question is likely to be less informative for White patients, who are typically well-represented in clinical trial evidence, we also asked respondents about how confident they are about the effectiveness of a

drug approved on the basis of evidence generated entirely outside of the United States. Such a scenario mirrors a recent trend in the "offshoring" of clinical trials.

For all respondents who were not highly confident about extrapolating — which turns out to be the vast majority — we sought to understand the rationale for their beliefs. In particular, we asked why they believed a drug on one sample would not perform equally well in another. Our multiple choice responses included worries that the drug would not work the same due to biological factors, socioeconomic and environmental factors, trust in the trial or other. Respondents who selected "other" provided open-text answers.

Results are reported in Table V. Panel (a) presents views from Black patients and doctors who treat them regarding extrapolating across race. Panel (b) presents views from White patients and doctors who treat them regarding confidence in extrapolating across geography. Each cell demonstrates the percentage of respondents who fall into that category. We find three broad patterns. First, few people fall into the highest confidence category for this exercise: the mean share of highly confident respondents across all four sets is 10.4 — ranging from 5.1 percent among PBPs to 19.6 percent among PWPs. Second, patients are less confident extrapolating on average than physicians: the mean level of confidence for Black and White patients is 1.0 (std. dev. 0.97) and 1.3 (std. dev. 0.91), respectively. For physicians treating these groups, the values are 1.72 (std. dev. 0.65) and 1.91 (std. dev. 0.65), respectively. In both instances, confidence among White patients and their doctors is slightly higher than their counterparts in Panel (b). Third, when providing a rationale for why a drug might work differently across samples, a nontrivial share select biological factors, though the most commonly chosen answer was socioeconomic and environmental factors.

Several doctors selected "other" and their open-text responses are reproduced in Appendix Table D1. When discussing extrapolation across race, doctors mention external validity, skepticism with results not obtained from representative samples, or a normative desire for the inclusion of diverse populations. With respect to foreign trial data, similar concerns were raised, though physicians also wondered about standards for studies performed abroad. One respondent noted that ease of extrapolation depends on where the study took place, stating: "It would depend upon the country. I would expect Western European and Canadian trials to be similar to my particular patient population."

Table V: Extrapolation from Clinical Trial Data among Physicians and Patients

| | Panel A: Black Patients and Physicians Who Treat Them (PBP) | | | | | |
|---|---|---|---|---|---|---|
| **White to Black Patients** | *Confidence* | | | | *Rationale* | |
| | Not at All | Some | Moderate | High | Perecieved Biol. Factors | Percieved Social & Envir. Factors |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Black Patients | 39.6% | 28.1% | 25.2% | 7.2% | 31.0% | 45.7% |
| PBP | 3.5% | 28.1% | 61.4% | 7.0% | 32.1% | 45.3% |
| | **Panel B: White Patients and Physicians Who Treat Them (PWP)** | | | | | |
| **Offshored to U.S. Patients** | *Confidence* | | | | *Rationale* | |
| | Not at All | Some | Moderate | High | Perecieved Biol. Factors | Percieved Social & Envir. Factors |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| White Patients | 21.3% | 36.8% | 32.4% | 9.6% | 19.5% | 43.9% |
| PWP | 1.5% | 21.5% | 61.5% | 15.4% | 10.9% | 70.9% |

*Notes:* Table reports clinical trial data extrapolation confidence and rationale among patients and physicians. Panel (a) reports confidence in extrapolation across race among *Black Patients* and *PBP*. Panel (b) reports confidence in extrapolation across geography among *White Patients* and *PWP*. Columns (1)–(4) report the percentage of respondents at each confidence level. If a respondent did not select "High" confidence in extrapolation, they were asked to provide a rationale. Column (5) reports the percentage of respondents who cite perceived biol. factors as the rationale for not having "High" confidence in extrapolation. Column (6) reports the percentage of respondents who cite perceived social and envir. factors as the rationale for not having "High" confidence in extrapolation. For each subgroup (*Black Patients, White Patients, PBP, PWP*), Appendix Table C11 reports confidence and rationale for both extrapolation questions (race and geography). *PBP* (Physicians treating Black patients) denotes physicians who report above the median percent Black patients in their patient panel. *PWP* (Physicians treating White patients) is defined similarly with respect to White patients.

### V.2.4 Threats to Internal Validity

Concerns with survey responses as outcomes include social desirability or experimenter demand effects. As mentioned above, we added consequentiality to both the physician (*i.e.*, reporting findings on trial preferences to federal agencies) and patient (*i.e.*, personalized reports to be shared with their doctors) experiments. The majority of physicians and nearly half of all patient respondents requested access to these reports, suggesting they were indeed of value to participants. For the patient survey, all respondents had been diagnosed with hypertension and thus had limited incentives to distort responses to information about a new drug of potential health benefit for their specific condition. Our results on subsamples of respondents who asked for the reports are similar to those reported above (see Appendix Tables C12 and C6 for experimental results from physicians and patients, respectively).[56]

The second key feature that reduces concerns about social desirability or experimenter demand effects is that we pre-specified heterogeneity by the race of the patient and the type of provider (those that treat more vs. fewer Black patients). If social desirability was playing a large role, patterns might be similar

[56]Appendix Tables C4 and C13 show that patients and doctors who downloaded or requested the report are statistically similar to other respondents. We also find that the order of profiles presented to physicians does not substantially impact their responsiveness to the treatment (Appendix Figure B13).

across Black and White patient respondents and across doctors treating all types of patients. In terms of experimenter demand, the patients were only shown one trial so it would have been difficult for them to discern the rationale for the study. Indeed, a word cloud of responses to the open-ended question "What do you think this study was about?" shows only limited references to race or diversity (see Appendix Figure B12) with the dominant response being "Blood Pressure." Similarly, our physician survey closely resembled the demographic information presented in biomedical publications and regulatory publications (*e.g.*, the FDA Drug Trial Snapshots database). Because we did not emphasize race or diversity beyond what physicians are typically exposed to, the extent of experimenter demand effects is plausibly limited (see Appendix Figure B14).

We follow Kuziemko et al. (2015) and Elías, Lacetera and Macis (2019), who use donations and petitions to validate survey responses and ask physicians to make a decision about a donation in a follow-up survey.[57] Our follow-up donation survey finds that the amount physicians allocate to the enrollment campaign targeting underrepresented minorities is strongly and significantly associated with physicians-specific coefficients on representation (Table VI Columns 1 and 2) and not with physician-specific responsiveness to efficacy. As the donation question was fielded to physicians as a follow-up question released 1–3 weeks after they completed the survey experiment, the results also suggest that our findings are unlikely to be driven by experimenter demand. We elect not to use willingness to pay, donations made by survey respondents, or petitions for our patient sample, as racial gaps in both the ability to pay and demand for privacy may complicate comparisons across groups (Alsan et al. 2020).

[57]We sent a follow-up survey after at least a week, in order to allow for some time between the actual survey and the donation question. There are few differences between our original sample and the sample of physicians who respond to the follow-up survey, with the exception of race. Physicians who reply to the donation question are more likely to be White than non-White (Appendix Table C4).

Table VI: Association Between Physician-Specific Coefficients and Trial Donations

| | (1) | (2) |
|---|---|---|
| Coefficient on Representation | 1.279*** | 1.229*** |
| | (0.449) | (0.436) |
| Coefficient on Efficacy | | 0.199 |
| | | (0.621) |
| Constant | 3.534 | 3.485 |
| Observations | 82 | 82 |

*Notes:* Table reports OLS estimates from a regression of physician-specific coefficients for representation and efficacy on dollars donated to a campaign to increase the representativeness of clinical trials. Physicians were asked to indicate, out of a possible $5, how many dollars they would like the research team to donate to a campaign that advocates for increases in clinical trial representation versus a campaign that advocates for increases in participation in clinical trials more generally. Observations are at the physician level. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

### V.2.5 Threats to External Validity

There are several potential and interrelated concerns related to mapping our survey results to real-world behavior. First is the potential issue that we prime people to think about something obviously bad, which can impact their survey responses. Second is the possibility that we induce patients to construct beliefs on-the-fly about something (clinical trials) on which they are not well informed. Third is that even if people do know about clinical trials and even if they are aware of their lack of representation, features of trials may not alter real-world decisions about prescribing and medication adherence.

Regarding the notion that we used an obviously negative prime (underrepresentation) for Black respondents, this presumes that ex-ante we had access to our ex-post results. Recall that our null hypothesis was that representation did *not* matter, which is precisely what we can now reject. Thus, we view our design as making underrepresentation — a widely known aspect of medical research — especially salient in the context of the survey experiment. We also ask an open-text question immediately after the intervention for our patient respondents about the rationale for their responses; sentiment analysis reported in Table C15 indicates no significant difference in positive affect across race groups. Further, the time spent on the survey does not differ across those groups.

Of course, if patients are unaware of clinical trials and our surveys elicit responses that do not map onto real behaviors, our findings are less relevant. However, data from Research!America and our own follow-up survey indicate that patients are, in fact, aware of clinical trials and that Black patients believe that they are not well represented in trial samples. Returning to the Research!America data in Table I,

Column (1) indicates that on average, 80 percent of Black respondents report that they have heard of clinical trials.

Regarding whether this information matters in practice, we document that it affects prescribing intention and updating. In the real world, we do see skepticism of research institutions and FDA-approved technologies (even those yet to be developed) among Black respondents in Table I. Qualitative comments from physicians in our study, as well as those attributed to a recent NASEM report, also suggest representation plays a role in how doctors practice medicine (see Appendix Tables D1 and D2).

### V.2.6 Robustness

We probe the robustness of our findings for physicians in Appendix Table C12. Columns (1) and (2) indicate that we obtain similar results when we use non-standardized versions of the outcomes. We replicate our main findings with standardized prescribing as the outcome in Column (3) and show that our findings are largely unchanged when restricting the sample to physicians who answer our follow-up donation question (Column 4) or among those who request a copy of our report to NIH and NASEM (Column 5). Column (6) shows findings on representation are not sensitive to the addition of controls selected using double-selection LASSO linear regression (Chernozhukov et al. 2018).

Additional results from our physician sample are presented in Appendix Table C14. Column (1) reports our main results from Equation 2, while Column (2) assesses whether representation and efficacy are substitutes or complements by adding an interaction term; we find no evidence of either.[58] Columns (4)–(6) indicate that our finding of substantial heterogeneity by Black patient representation in one's panel is insensitive to varying definitions of physicians who treat Black patients. Our finding of a strong interaction between representation and reported patient percent Black (from Table III and replicated in Column 3) is robust to dichotomizing patient percent Black at the median as well as to defining physicians treating Black patients using zip code-level statistics obtained from the U.S. Census Bureau. In Appendix Figure B15, we present further tests of robustness, including results from alternative specifications and on the sample of observations with at least one efficacy duplicate, and show that our finding of a significant coefficient on representation withstands all these tests.

We report robustness checks for our patient experiment in Appendix Table C6. Panel (a) demonstrates that results across our three outcomes are unchanged when we restrict to patients who requested the personalized report we offered, whereas Panel (b) shows that our findings are robust to weighting patients using person weights obtained from the Medical Expenditure Panel Survey. Panel (c) indicates that our results are robust to including LASSO-selected controls.

[58] See Appendix A.2 for additional discussion.

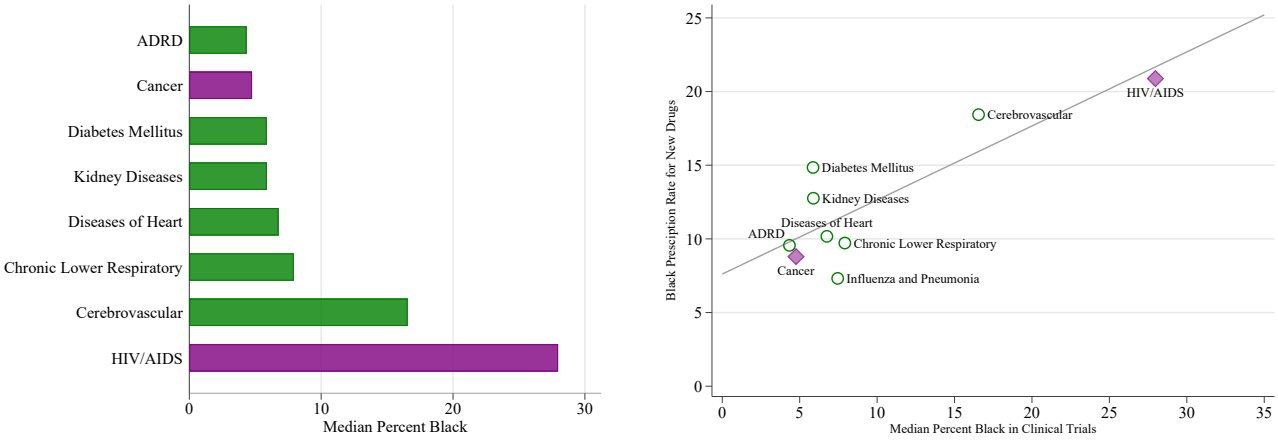
# VI Discussion and Conclusion

## VI.1 Inclusive Infrastructure: A Case Study

Recall that Section III detailed a key market failure that helps to explain the current lack of representation in medical evidence: "missing" investments in *inclusive infrastructure*, which arise when no firm fully internalizes the benefits of developing inclusive recruitment systems. Here, we combine quantitative and qualitative evidence — including interviews with experts in trial design — to tighten the links between our theoretical and empirical findings and real-world practice.

Figure VI Panel (a) plots the median percent Black in pivotal trials across the most common diseases or conditions in the United States.[59] Black patients are underrepresented relative to their population share across most conditions, and underrepresented relative to disease burden as well (see Appendix Figure B16), though there is significant variation across conditions. In Panel (b), we document that higher representation of Black patients in clinical trials is associated with higher outpatient prescribing of new drugs that fall into that disease category among Black patients.

Figure VI: Trial Representation by Condition and Association with New Drug Prescribing




(a) Median Percent Black in Trials

(b) Prescription Rates and Trial Representation

*Notes:* Panel (a) plots the median share of Black patients in trials across HIV/AIDS and the ten leading causes of death (excluding unintentional injuries and suicide) in the United States (Heron 2021). Data on trial composition are drawn from ClinicalTrials.gov. Panel (b) plots the correlation between the prescription rate of new medications to Black Americans and the median percent Black in pivotal trials. We construct the prescription rate as the percentage of newly marketed drugs (on the market for five or fewer years) received by Black Americans in each major condition category. In Panel (b), the y-axis value of "Cancer" refers to outpatient cancer supportive therapies. Prescription data are drawn from the Medical Expenditure Panel Survey. Observations associated with cancer and HIV/AIDS are marked in diamonds (purple). We include a line of best fit. See Appendix H.1.1 and H.2 for details on data.

Next, we focus on cancer and HIV/AIDS (purple diamonds in Figure VI Panel b), which are instructive to compare for several reasons. Both disease areas benefit from decades of federal investments into

[59] All diseases or conditions presented except HIV/AIDS are among the ten leading causes of death in the United States (Heron 2021). We did not include unintentional injuries and suicide as there are no pharmaceuticals intended to prevent/treat such deaths.

research networks across the U.S. by the National Cancer Institute (NCI) and National Institute of Allergy and Infectious Diseases (NIAID), respectively.[60] Federal investments into these networks are comparable, totaling $6.54 billion into NCI and $6.05 billion into NIAID in 2021 (Congressional Research Service 2022). Yet these networks have different origins, which shed light on differences in contemporary outcomes. While investment in cancer research was driven by top-down investments into academic medical centers, including the National Cancer Act of 1971 (Mukherjee 2010), HIV/AIDS research has been defined by active community involvement that fundamentally influences protocol design and recruitment methods.[61] Interviews with the HIV Vaccine Trial Network highlighted a key resulting difference: in contrast to cancer, HIV/AIDS research sites have historically been selected in conversation with community partners and thus are not limited to large academic medical centers.

Table VII substantiates these anecdotes and makes clear how site selection shapes trial composition. Amongst 273,842 U.S.-based trial sites listed in the ClinicalTrials.gov database, sites that enroll for HIV/AIDS are approximately 25 percentage points more likely to be located at a safety net hospital than sites that recruit for cancer or ADRD.[62] Unsurprisingly, the demographic characteristics of the trial sites also differ. Appendix Tables C17 and C18 report information on the demographics of HIV/AIDS, cancer, and ADRD research centers at the Hospital Service Area (HSA) level for all clinical trials and specific networks. Trial sites recruiting for cancer have, on average, a 10.4 percentage point higher share of non-Hispanic White population and a 3.0 percentage point higher share of those with private health insurance than trial sites recruiting for HIV/AIDS.[63]

Site selection is just one part of the R&D process — protocol development is another important step and also differs across conditions. Since 1990, The DAIDS at the National Institute of Allergy and Infectious Disease (NIAID) has required that trial protocols include explicit community engagement plans, developed in conjunction with standing community advisory boards (CABs) (Strauss et al. 2001).[64] CABs meet regularly with trial investigators and consult on proposed protocols. Our discussion with HIV Vaccine Trial Network (HVTN) leadership suggests DAIDS requirements have important spillover effects: although firms are not obligated to comply, industry sponsors often engage with communities in order to benefit from existing recruitment networks.

The stark differences in trial composition for cancer and HIV/AIDS highlight the extent to which

[60]There are 131 dedicated research centers that co-organize trials for cancer, and 108 co-organize trials for HIV/AIDS. Although the majority of HIV/AIDS funding is allocated via NIAID, the NCI also includes budgets for HIV/AIDS research.

[61]See Epstein (1996) for an excellent discussion of how activists shaped NIH-sponsored research and affected the development of clinical trials.

[62]As no formal definition of safety net hospitals exists, we proxy for it using the Medicaid Disproportionate Share Hospital (DSH) Index and uncompensated care as a percentage of total operating expenses (Popescu et al. 2019).

[63]Appendix Figure B17 demonstrates a strong correlation between trial site zip code share Black and share Black in a trial. See Appendix Section G.1 for information on recent cancer and ADRD initiatives to diversify site selection. One consideration is that quality might be inferior in safety net hospitals — we outline efforts to compensate patients for participation as well as improve the quality of hospitals that serve Black patients in Appendix Section G.1 (see also Chandra, Kakani and Sacarny 2020 for evidence of recent quality improvement in hospitals).

[64]Although many institutions maintain CABs for cancer trials, the CAB requirement at DAIDS is unique. See details on DAIDS requirements here: https://www.niaid.nih.gov/daids-ctu/community-engagement-NEW.

Table VII: Trial Sites at Safety Net Hospitals

| | DSH Index | | UCMP Care | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| HIV/AIDS (Cancer Comparison) | 0.251*** | | 0.035*** | |
| | (0.007) | | (0.006) | |
| HIV/AIDS (ADRD Comparison) | | 0.234*** | | 0.017* |
| | | (0.011) | | (0.009) |
| Constant | 0.393*** | 0.410*** | 0.142*** | 0.161*** |
| | (0.001) | (0.009) | (0.001) | (0.007) |
| Observations | 201,432 | 7,680 | 184,637 | 6,733 |

*Notes:* Table reports OLS estimates from a regression of an indicator for whether a trial site is at a safety net hospital (SNH). Each observation represents a specific site associated with a unique clinical trial and the data are limited to Cancer, HIV/AIDS, and ADRD trials. Following Popescu et al. (2019), we define a SNH as a hospital in the top quartile within the state it is located in either receiving a disproportionate share of funding from Medicaid or uncompensated care. In Columns (1) and (3), we use the Disproportionate Share Hospital (DSH) Index to define a SNH — these data are from CMS Cost Reports. In Columns (2) and (4), we use the cost of uncompensated care (as a percentage of total operating expenses) — these data are from the American Hospital Association's 2020 survey. *HIV/AIDS (Cancer Comparison)* is an indicator variable equal to one if a trial site studies HIV/AIDS and zero if a trial site studies cancer. *HIV/AIDS (ADRD Comparison)* is an indicator variable equal to one if a trial site studies HIV/AIDS and zero if a trial site studies ADRD. Trial site information is drawn from ClinicalTrials.gov. See Appendix H.1.1 and H.3.8 for details on data. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

active, large-scale investments in inclusive infrastructure, in addition to incentives, can be important for improving health equity. Figure VI Panel (b) demonstrates a positive relationship between better representation in trials and prescribing rates.[65] This descriptive finding is robust to dropping HIV/AIDS, (see Appendix Figure B18), though the main takeaway from this section is that HIV/AIDS *is* an "outlier" on many dimensions and therefore a useful template for reform.

## VI.2 Concluding Comments

This article investigates the causes and consequences of routine exclusion from the R&D process. Motivated by persistent, large disparities in both clinical trial participation and health outcomes, our empirical focus is on the underrepresentation of Black Americans in medical research. A theoretical model of similarity-based extrapolation clarifies the extent to which this exclusion shapes both patient and physician beliefs and behaviors. Black patients, and the physicians who treat them, find trial evidence less relevant for their care and are less likely to prescribe medications when experimental samples are not representative. However, when the evidence base is more racially representative, these updating and prescribing gaps close.

We present case studies on clinical trials associated with cancer and HIV/AIDS – disease areas that

[65] Another way HIV/AIDS is unique is Ryan White Care Act funding (see Dillender 2022). Title I funds cities and Title II funds states, a portion of which must go to the AIDS Drug Assistance Programs, which in turn, may have pull incentives on innovation as per Acemoglu et al. (2006), Finkelstein (2004), and Acemoglu and Linn (2004).

have benefited from decades of public sector investment, though have very different approaches to producing science. Interviews, qualitative evidence, and administrative data suggest that recruiting a racially diverse sample is in fact feasible but requires investments in inclusive infrastructure.

# References


**Acemoglu, Daron, and Joshua Linn.** 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *The Quarterly Journal of Economics*, 119(3): 1049–1090.

**Acemoglu, Daron, David Cutler, Amy Finkelstein, and Joshua Linn.** 2006. "Did Medicare Induce Pharmaceutical Innovation?" *American Economic Review*, 96(2): 103–107.

**Agha, Leila, and David Molitor.** 2018. "The Local Influence of Pioneer Investigators on Technology Adoption: Evidence from New Cancer Drugs." *Review of Economics and Statistics*, 100(1): 29–44.

**Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hémous.** 2019. "Innovation and Top Income Inequality." *Review of Economic Studies*, 86(1): 1–45.

**Alesina, Alberto, Armando Miano, and Stefanie Stantcheva.** 2022. "Immigration and Redistribution." *The Review of Economic Studies*.

**Alesina, Alberto, Edward Glaeser, and Bruce Sacerdote.** 2001. "Why Doesn't the United States Have a European-Style Welfare State?" *Brookings Papers on Economic Activity*, 2(1): 187–277.

**Allcott, Hunt.** 2015. "Site Selection Bias in Program Evaluation." *The Quarterly Journal of Economics*, 130(3): 1117–1165.

**Alsan, Marcella, and Marianne Wanamaker.** 2018. "Tuskegee and the Health of Black Men." *Quarterly Journal of Economics*, 133(1): 407–455.

**Alsan, Marcella, Luca Braghieri, Sarah Eichmeyer, Minjeong Joyce Kim, Stefanie Stantcheva, and David Y. Yang.** 2020. "Civil Liberties in Times of Crisis." *NBER Working Paper No. 27972*.

**Alsan, Marcella, Owen Garrick, and Grant Graziani.** 2019. "Does Diversity Matter for Health? Experimental Evidence from Oakland." *American Economic Review*, 109(12): 4071–4111.

**Alsan, Marcella, Vincenzo Atella, Jay Bhattacharya, Valentina Conti, Iván Mejía-Guevara, and Grant Miller.** 2019. "Technological Progress and Health Convergence: The Case of Penicillin in Post-War Italy." *NBER Working Paper No. 25541*.

**American Academy of Family Physicians.** 2022. "AAFP Membership FAQ." *https://www.aafp.org/membership/faq.html*, Accessed: 10/04/2022.

**American Board of Internal Medicine.** 2022. "Maintaining Certification Requirements." *https://www.abim.org/maintenance-of-certification/moc-requirements*, Accessed: 10/04/2022.

**American Diabetes Association.** 2020. "Section 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2021." *Diabetes Care*, 43: S111–S124.

**Anderson, Monique L, Karen Chiswell, Eric D Peterson, Asba Tasneem, James Topping, and Robert M Califf.** 2015. "Compliance with Results Reporting at ClinicalTrials.gov." *New England Journal of Medicine*, 372(11): 1031–1039.

**Andreoni, James.** 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-glow Giving." *The Economic Journal*, 100(401): 464–477.

**Arias, Elizabeth, Betzaida Tejada-Vera, Kenneth D. Kochanek, and Farida B. Ahmad.** 2022. "Provisional Life Expectancy Estimates for 2021." 23, American Medical Association.

**Ayres, Ian.** 2010. "Testing for Discrimination and the Problem of Included Variable Bias." *Yale Law School Mimeo*.

**Bach, Peter B., Hoangmai H. Pham, Deborah Schrag, Ramsey C. Tate, and J. Lee Hargraves.** 2004. "Primary Care Physicians Who Treat Blacks and Whites." *New England Journal of Medicine*, 351(6): 575–584.

**Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Lawrence F. Katz, Benjamin A. Olken, and Anja Sautmann.** 2020. "In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plans for RCTs in Economics." *NBER Working Paper No. 26993*, 1–14.

**Barnato, Amber E.** 2017. "Challenges In Understanding And Respecting Patients' Preferences." *Health Affairs*, 36(7): 1252–1257.

**Basu, Anirban, and Kritee Gujral.** 2020. "Evidence Generation, Decision making, and Consequent Growth in Health Disparities." *Proceedings of the National Academy of Sciences*, 117(25): 14042–14051.

**Blanco, Maria A., Josephine L. Dorsch, Gerald Perry, and Mary L. Zanetti.** 2014. "A Survey Study of Evidence-Based Medicine Training in US and Canadian Medical Schools." *Journal of the Medical Library Association*, 102(3): 160–168.

**Blewett, Lynn A., Julia A. Rivera Drew, Risa Griffin, and Kari C.W. Williams.** 2019. "IPUMS Health Surveys: Medical Expenditure Panel Survey, Version 1.1 [2019]." *https://doi.org/10.18128/D071.V1.1*.

**Bordalo, Pedro, B Giovanni, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer.** 2022. "Imagining the Future: Memory, Simulation and Beliefs About COVID." *NBER Working Paper No. 30353*.

**Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer.** 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

**Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer.** 2019. "Beliefs About Gender." *American Economic Review*, 109(3): 739–73.

**Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer.** 2020. "Memory, Attention, and Choice." *Quarterly Journal of Economics*, 135(3): 1399–1442.

**Budish, Eric, Benjamin N. Roin, and Heidi Williams.** 2015. "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials." *American Economic Review*, 105(7): 2044–2085.

**Center for Disease Control and Prevention.** 2021. "Prevalence of Both Diagnosed and Undiagnosed Diabetes." *https://www.cdc.gov/diabetes/data/statistics-report/diagnosed-undiagnosed-diabetes.html*, Accessed: 10/04/2022.

**Centers for Disease Control and Prevention.** 2021. "HIV Surveillance Report, 2019." Accessed: 10/04/2022.

**Centers for Medicare & Medicaid Services.** 2019. "Healthcare Cost Report Information System." Accessed through the American Hospital Association.

**Centers for Medicare  Medicaid Services.** 2022. "Disproportionate Share Hospital (DSH)." *âĂŒhttps://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/dsh*, Accessed: 10/04/2022.

**Chandra, Amitabh, and Jonathan Skinner.** 2003. "Geography and Racial Health Disparities." *NBER Working Paper No. 9513*.

**Chandra, Amitabh, Michael Frakes, and Anup Malani.** 2017. "Challenges to Reducing Discrimination and Health Inequity Through Existing Civil Rights Laws." *Health Affairs*, 36(6): 1041–1047.

**Chandra, Amitabh, Pragya Kakani, and Adam Sacarny.** 2020. "Hospital Allocation and Racial Disparities in Health Care." *NBER Working Paper No. 28018*.

**Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins.** 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *The Econometrics Journal*, 21(1): C1–C68.

**Cochrane, Archie.** 1972. *Effectiveness and Efficiency: Random Reflections on Health Services.* London, England:The Nuffield Provincial Hospital Trust.

**Congressional Research Service.** 2022. "National Institutes of Health (NIH) Funding: FY1996-FY2023." *https://sgp.fas.org/crs/misc/R43341.pdf* , Accessed: 10/04/2022.

**Cutler, David M., Ellen Meara, and Seth Richards-Shubik.** 2012. "Induced Innovation and Social Inequality: Evidence from Infant Medical Care." *Journal of Human Resources*, 47(2): 456–492.

**Darity, William, Jr., A. Kirsten Mullen, and Marvin Slaughter.** 2022. "The Cumulative Costs of Racism and the Bill for Black Reparations." *Journal of Economic Perspectives*, 36(2): 99–122.

**Daw, Nathaniel D.** 2014. "Advanced Reinforcement Learning." *Neuroeconomics*, 299–320.

**Dillender, Marcus.** 2022. "Evidence and Lessons on the Health Impacts of Public Health Funding from the Fight Against HIV/AIDS." *NBER Working Paper No. 28867*.

**DiMasia, Joseph A., Henry G. Grabowski, and Ronald W. Hansen.** 2016. "Innovation in the Pharmaceutical Industry: New Estimates of RD Costs ." *Journal of Health Economics*, 47: 20–33.

**Ding, Dong, and Sherry A. Glied.** 2022. "Disparities in the Use of New Diabetes Medications: Widening Treatment Inequality by Race and Insurance Coverage." *Commonwealth Fund*.

**Dornsife, Dana, Stephanie Monroe, Nicole Richie, Fabian Sandoval, Claudine Brisard, Nicholas Kenny, Keri McDonough, and Kathleen Starr.** 2019. "How to Boost Racial, Ethnic and Gender Diversity in Clinical Research." *Syneos Health*.

**Doximity.** 2014. "Survey: How Doctors Read and What it Means to Patients." *https://www.businesswire.com/news/home/20140722005535/en/Survey-Doctors-Read-Means-Patients*, Accessed: 10/04/2022.

**Ehrhardt, Stephan, Lawrence J. Appel, and Curtis L. Meinert.** 2015. "Trends in National Institutes of Health Funding for Clinical Trials Registered in ClinicalTrials.gov." *JAMA*, 314(23): 2566–2567.

**Einav, Liran, Amy Finkelstein, Tamar Oostrom, Abigail Ostriker, and Heidi Williams.** 2020. "Screening and Selection: The Case of Mammograms." *American Economic Review*, 110(12): 3836–70.

**Elhussein, Ahmed, Andrea Anderson, Michael P Bancks, Mace Coday, William C Knowle, Anne Peters, Elizabeth M Vaughan, Nisa M. Maruthur, Jeanne M Clark, and Scott Pilla.** 2022*a*. "Racial-Ethnic Disparities in Uptake of New Hepatitis C Drugs in Medicare." *Lancet Regional Health - Americas*, 6(100111): 1147–1158.

**Elhussein, Ahmed, Andrea Anderson, Michael P Bancks, Mace Coday, William C Knowler, Anne Peters, Elizabeth M Vaughan, Nisa M. Maruthur, Jeanne M Clark, and Scott Pilla.** 2022*b*. "Racial/Ethnic and Socioeconomic Disparities in the Use of Newer Diabetes Medications in the Look AHEAD study." *The Lancet Regional Health - Americas*, 6: 100111.

**Eli, Shari, Trevon D Logan, and Boriana Miloucheva.** 2019. "Physician Bias and Racial Disparities in Health: Evidence from Veterans' Pensions." *NBER Working Paper No. 25846*.

**Ellis, Randall P., and Thomas G. McGuire.** 1986. "Provider Behavior Under Prospective Reimbursement: Cost Sharing and Supply." *Journal of Health Economics*, 5(2): 129–151.

**Elías, Julio J., Nicola Lacetera, and Mario Macis.** 2019. "Paying for Kidneys? A Randomized Survey and Choice Experiment." *American Economic Review*, 109(8): 2855–2888.

**Epstein, Steven.** 1996. *Impure Science: AIDS, Activism, and the Politics of Knowledge.* Berkeley, CA:University of California Press.

**Epstein, Steven.** 2007. *Inclusion: The Politics of Difference in Medical Research.* Chicago, IL:University of Chicago Press.

**Federation of State Medical Boards.** 2022. "About Physician Licensure: How Physicians Gain Licenses to Practice Medicine." *https://www.fsmb.org/u.s.-medical-regulatory-trends-and-actions/guide-to-medical-regulation-in-the-united-states/about-physician-licensure*, Accessed: 10/04/2022.

**Fehr, Ernst.** 2009. "On the Economics and Biology of Trust." *Journal of the European Economic Association*, 7(2): 235–266.

**Feldman, Sergey, Waleed Ammar, Kyle Lo, Elly Trepman, Madeleine van Zuylen, and Oren Etzioni.** 2019. "Quantifying sex bias in clinical studies at scale with automated data extraction." *JAMA Network Open*, 2(7): e196700–e196700.

**Finkelstein, Amy.** 2004. "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry." *The Quarterly Journal of Economics*, 119(2): 527–564.

**Food and Drug Administration.** 2020. "Civil Money Penalties Relating to the ClinicalTrials.gov Data Bank." *https://www.fda.gov/media/113361/download*, Accessed: 10/04/2022.

**Food and Drug Administration.** 2022. "FDA Code of Federal Regulations Title 21." *https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRsearch.cfm?CFRPart=312*, Accessed: 10/04/2022.

**Foster, Andrew D, and Mark R Rosenzweig.** 2010. "Microeconomics of Technology Adoption." *Annual Review of Economics*, 2(1): 395–424.

**George, Sheba, Nelida Duran, and Keith Norris.** 2014. "A Systematic Review of Barriers and Facilitators to Minority Research Participation Among African Americans, Latinos, Asian Americans, and Pacific Islanders." *American Journal of Public Health*, 104(2): e16–31.

**Gilboa, Itzhak, and David Schmeidler.** 1995. "Case-Based Decision Theory." *The Quarterly Journal of Economics*, 110(3): 605–639.

**Glaeser, Edward L., David I. Laibson, José A. Scheinkman, and Christine L. Soutter.** 2000. "Measuring Trust." *Quarterly Journal of Economics*, 115(3): 811–846.

**Glied, Sherry, and Adriana Lleras-Muney.** 2008. "Technological Innovation and Inequality in Health." *Demography*, 45(3): 741–761.

**Golden, Sherita Hill, Arleen Brown, Jane A Cauley, Marshall H Chin, Tiffany L Gary-Webb, Catherine Kim, Julie Ann Sosa, Anne E Sumner, and Blair Anton.** 2012. "Health disparities in Endocrine Disorders: Biological, Clinical, and Nonclinical Factors-An Endocrine Society Scientific Statement." *Journal of Clinical Endocrinology and Metabolism*, 97(9): E1579–639.

**Goldman, Dana, and Darius Lakdawalla.** 2018. "The Global Burden of Medical Innovation." *USC Schaeffer Center for Health Policy & Economics*, , (White Paper).

**Green, Angela K., Niti Trivedi, Jennifer J. Hsu, Nancy L. Yu, Peter B. Bach, and Susan Chimonas.** 2022. "Despite The FDA's Five-Year Plan, Black Patients Remain Inadequately Represented In Clinical Trials For Drugs." *Health Affairs*, 41(3): 368–374.

**Greenwood, Brad N., Rachel R. Hardeman, Laura Huang, and Aaron Sojourner.** 2020. "Physician–patient Racial Concordance and Disparities in Birthing Mortality for Newborns." *Proceedings of the National Academy of Sciences (PNAS)*, 117(35): 21194–21200.

**Gruber, Jonathan, and Michael D. Frakes.** 2022. "Racial Concordance and the Quality of Medical Care: Evidence from the Military." *NBER Summer Institute 2022*.

**Gupta, Harsh.** 2022. "Do Female Researchers Increase Female Enrollment in Clinical Trials?" *Working paper*.

**Haaland, Ingar, Christopher Roth, and Johannes Wohlfart.** Forthcoming. "Designing Information Provision Experiments." *Journal of Economic Literature*.

**Hamilton, Barton H., Andrés Hincapié, Emma C. Kalish, and Nicholas W. Papageorge.** 2021. "Innovation and Health Disparities During an Epidemic: The Case of HIV." *NBER Working Paper No. 28864*.

**Heron, Melonie.** 2021. "Deaths: Leading Causes for 2019." *National Vital Statistics Reports*, 70(9).

**Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini.** 2021. "How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities." *American Economic Review*, 111: 1442–80.

**Hoffman, Kelly M, Sophie Trawalter, Jordan R Axt, and M Norman Oliver.** 2016. "Racial bias in Pain Assessment and Treatment Recommendations, and False Beliefs about Biological Differences between Blacks and Whites." *Proceedings of the National Academy of Sciences*, 113(16): 4296–4301.

**ICMJE.** 2022. "Clinical Trials Registration." *https://www.icmje.org/about-icmje/faqs/clinical-trials-registration/*, Accessed: 10/04/2022.

**IQVIA.** 2015. "Global Medicines Use in 2020: Outlook and Implications." IMS Institute for Healthcare Informatics Technical Report.

**Ivarsson, Melanie.** 2021. "Diversifying Clinical Trials and Clinical Research: The Industry Perspective." *https://www.nationalacademies.org/event/03-29-2021/overcoming-barriers-to-diversifying-clinical-trial-workshop*.

**Jaravel, Xavier.** 2019. "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector." *Quarterly Journal of Economics*, 134(2): 715–783.

**Jehiel, Philippe.** 2005. "Analogy-Based Expectation Equilibrium." *Journal of Economic Theory*, 123(2): 81–104.

**Jensen, Robert.** 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics*, 125(2): 515–548.

**Jones, Charles I., and Jihee Kim.** 2018. "A Schumpeterian Model of Top Income Inequality." *Journal of Political Economy*, 126(5): 1785–1826.

**Jung, Jeah, and Roger Feldman.** 2017. "Racial-Ethnic Disparities in Uptake of New Hepatitis C Drugs in Medicare." *Journal of Racial and Ethnic Health Disparities*, 4(6): 1147–1158.

**Kanich, Chris, Christian Kreibich, Kirill Levchenko, Brandon Enright, Geoffrey M. Voelker, Vern Paxson, and Stefan Savage.** 2009. "Spamalytics: An Empirical Analysis of Spam Marketing Conversion." *Communications of the ACM*, 52(9): 99–107.

**Kesselheim, Aaron S., Christopher T. Robertson, Jessica A. Myers, Susannah L. Rose, Victoria Gillet, Kathryn M. Ross, Robert J. Glynn, Steven Joffe, and Jerry Avorn.** 2012. "A Randomized Study of How Physicians Interpret Research Funding Disclosures." *New England Journal of Medicine*, 367(12): 1119–1127.

**Kline, Patrick, Neviana Petkova, Heidi L. Williams, and Owen Zidar.** 2019. "Who Profits from Patents? Rent-Sharing at Innovative Firms." *Quarterly Journal of Economics*, 134(3): 1343–1303.

**Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz.** 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83–119.

**Knepper, Todd C., and Howard L. McLeod.** 2018. "When Will Clinical Trials Finally Reflect Diversity?" *Nature*, 557: 157–159.

**Koning, Rembrand, Sampsa Samila, and John-Paul Ferguson.** 2021. "Who Do We Invent For? Patents by Women Focus More on Women's health, but Few Women Get to Invent." *Science*, 372(6548): 1345–1348.

**Kowalski, Amanda E.** 2022. "Behavior within a Clinical Trial and Implications for Mammography Guidelines." *The Review of Economic Studies*.

**Kremer, Michael, and Rachel Glennerster.** 2004. *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases.* Princeton, NJ:Princeton University Press.

**Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva.** 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review*, 105(4): 1478–1508.

**Ledley, Fred D., Sarah Shonka McCoy, Gregory Vaughan, and Ekaterina Galkina Cleary.** 2020. "Profitability of Large Pharmaceutical Companies Compared With Other Large Public Companies." *JAMA*, 323(9): 834–843.

**Logan, Trevon D, and Samuel L Myers Jr.** 2022. "Symposium: Race and Economic Literature – Introduction." *Journal of Economic Literature*, 60(2): 375–76.

**Malmendier, Ulrike, and Laura Veldkamp.** 2022. "Information Resonance." *Working Paper*.

**Martí-Carvajal, Arturo.** 2020. "What is Evidence-Based Medicine?" *Ciencias Médicas*, 1(23): 1–7.

**Masic, Izet, Milan Miokovic, and Belma Muhamedagic.** 2008. "Evidence Based Medicine - New Approaches and Challenges." *Acta Informatica Medica*, 16(4): 219–225.

**McCoy, Rozalina G., Hayley J. Dykhoff, Lindsey Sangaralingham, Joseph S. Ross, Pinar Karaca-Mandic, Victor M. Montori, and Nilay D. Shah.** 2019. "Adoption of New Glucose-Lowering Medications in the U.S.—The Case of SGLT2 Inhibitors: Nationwide Cohort Study." *Diabetes Technology and Therapeutics*, 21(12): 702–712.

**Michelman, Valerie, and Lucy Msall.** 2021. "Sex, Drugs, and RD: Missing Innovation from Regulating Female Enrollment in Clinical Trials." *Working paper*.

**Moderna Therapeutics.** 2020. "2020 Cove Study Enrollment Completion - 10/22/2020." *https://www.modernatx.com/sites/default/files/content_documents/2020-COVE-Study-Enrollment-Completion-10.22.20.pdf*, Accessed: 02/17/2022.

**Moore, Thomas J., Hanzhe Zhang, Gerard Anderson, and G. Caleb Alexander.** 2018. "Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration, 2015-2016 ." *JAMA Internal Medicine*, 178(11): 1451–1457.

**Mukherjee, Siddhartha.** 2010. *The Emperor of All Maladies: A Biography of Cancer.* New York, NY:Simon and Schuster.

**Mullainathan, Sendhil.** 2002. "A Memory-Based Model of Bounded Rationality." *The Quarterly Journal of Economics*, 117(3): 735–774.

**Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer.** 2008. "Coarse Thinking and Persuasion." *The Quarterly Journal of Economics*, 123(2): 577–619.

**NASEM.** 2022. *Improving Representation in Clinical Trials and Research: Building Research Equity for Women and Underrepresented Groups.* Washington, DC:National Academies Press.

**Nathan, David M., John B. Buse, Mayer B. Davidson, Ele Ferrannini, Rury R. Holman, Robert Sherwin, and Bernard Zinman.** 2009. "Medical Management of Hyperglycemia in Type 2 Diabetes: A Consensus Algorithm for the Initiation and Adjustment of Therapy: A Consensus Statement of the American Diabetes Association and the European Association for the Study of Diabetes." *Diabetes Care*, 32(1): 193–203.

**National Institutes of Health.** 2017. "NIH's Definition of a Clinical Trial." *https://grants.nih.gov/policy/clinical-trials/definition.htm*.

**National Institutes of Health.** 2020. "Phase 3 Clinical Trial of Investigational Vaccine for COVID-19 Begins." *https://www.nih.gov/news-events/news-releases/phase-3-clinical-trial-investigational-vaccine-covid-19-begins*, Accessed: 10/04/2022.

**Ornstein, Charles, Tracy Weber, and Ryann Grochowski Jones.** 2019. "We Found Over 700 Doctors Who Were Paid More Than a Million Dollars by Drug and Medical Device Companies." *https://www.propublica.org/article/we-found-over-700-doctors-who-were-paid-more-than-a-million-dollars-by-drug-and-medical-device-companies*, Accessed: 10/04/2022.

**Ostchega, Yechiam, Cheryl D. Fryar, Tatiana Nwankwo, and D.O. Duong T. Nguyen.** 2020. "Hypertension Prevalence Among Adults Aged 18 and Over: United States, 2017–2018." *National Center for Health Statistics Data Brief*, , (364): 1–7.

**Oster, Emily.** 2020. "Health Recommendations and Selection in Health Behaviors." *American Economic Review: Insights*, 2(2): 143–60.

**Papageorge, Nicholas W.** 2016. "Why Medical Innovation is Valuable: Health, Human Capital and the Labor Market." *Quantitative Economics*, 7(3): 671–725.

**Popescu, Ioana, Kathryn R. Fingar, Eli Cutler, Jing Guo, and H. Joanna Jiang.** 2019. "Comparison of 3 Safety-Net Hospital Definitions and Association With Hospital Characteristics." *JAMA Network Open*, 2(8): e198577–e198577.

**Qiao, Yao, G. Caleb Alexander, and Thomas J. Moore.** 2019. "Globalization of Clinical Trials: Variation in Estimated Regional Costs of Pivotal Trials, 2015-2016." *Clinical Trials*, 16(3): 329–333.

**Richardson, Matthew, Ewa Dominowska, and Robert Ragno.** 2007. "Predicting clicks: Estimating the Click-through Rate for New Ads." *Proceedings of the 16th International Conference on the World Wide Web*, 521–530.

**Roberts, Dorothy.** 2012. *Fatal Invention: How Science, Politics, and Big Business Re-create Race in the Twenty-first Century.* New York, NY:The New Press.

**Roth, Christopher, and Johannes Wohlfart.** 2020. "How Do Expectations about the Macroeconomy Affect Personal Expectations and Behavior?" *The Review of Economics and Statistics*, 102(4): 731–748.

**Roth, Jean.** 2018. "NDC to Labeler Code Product Code Package Size Crosswalk." *National Bureau of Economic Research*. Accessed: 10/04/2022.

**Sapienza, Paola, Anna Toldra, and Luigi Zingales.** 2013. "Understanding Trust." *Economic Journal, Royal Economic Society*, 123(12).

**Schwartz, Lisa M., and Steven Woloshin.** 2019. "Medical Marketing in the United States, 1997-2016." *JAMA*, 321(1): 80–96.

**Sertkaya, Aylin, Anna Birkenbach, Ayesha Berlind, and John Eyraud.** 2014. "Examination of Clinical Trial Costs and Barriers for Drug Development Final." *https://aspe.hhs.gov/reports/examination-clinical-trial-costs-barriers-drug-development-0*.

**Skinner, Jonathan, and Douglas Staiger.** 2005. "Technology Adoption from Hybrid Corn to Beta Blockers." *NBER Working Paper No. 11251*.

**Skinner, Jonathan, and Douglas Staiger.** 2015. "Technology Diffusion and Productivity Growth in Health Care." *Review of Economics and Statistics*, 97(5): 951–964.

**Smith, Tom W., Michael Davern, Jeremy Freese, and Stephen L. Morgan.** 2021. "General Social Surveys, 1972-2021: Cumulative Codebook."

**Sosinsky, Alexandra Z., Janet W. Rich-Edwards, Aleta Wiley, Kalifa Wright, Primavera A. Spagnolo, and Hadine Joffe.** 2022. "Enrollment of Female participants in United States Drug and Device Phase 1–3 Clinical Trials Between 2016 and 2019." *Contemporary Clinical Trials*, 115.

**Stantcheva, Stefanie.** 2022. "Understanding Tax Policy: How do People Reason." *Quarterly Journal of Economics*, 136(4): 2309–2369.

**STAT News.** 2020. *Representation and Diversity in Clinical Trials.* Boston, MA.

**Steinberg, Jecca R, Brandon E Turner, Brannon T Weeks, Christopher J Magnani, Bonnie O Wong, Fatima Rodriguez, Lynn M Yee, and Mark R Cullen.** 2021. "Analysis of Female Enrollment and Participant Sex by Burden of Disease in US clinical Trials Between 2000 and 2020." *JAMA Network Open*, 4(6): e2113749–e2113749.

**Strauss, R. P., S. Sengupta, S C Quinn, C. Spaulding J. Goeppinger, S. M. Kegeles, and G. Millett.** 2001. "The Role of Community Advisory Boards: Involving Communities in the Informed Consent Process." *American Journal of Public Health*, 91(12): 1938–1943.

**Tasneem, Asba, Laura Aberle, Hari Ananth, Swati Chakraborty, Karen Chiswell, Brian J. McCourt, and Ricardo Pietrobon.** 2012. "The Database for Aggregate Analysis of ClinicalTrials.gov (AACT) and Subsequent Regrouping by Clinical Specialty." *PLoS ONE*, 7(3): e33677.

**Tirrell, Meg, and Leanne Miller.** 2020. "Moderna Slows Coronavirus Vaccine Trial Enrollment to Ensure Minority Representation, CEO Says." *CNBC*.

**U.S. Census Bureau.** 2021. "Census Bureau QuickFacts." *https://www.census.gov/quickfacts/fact/table/US/PST045221*, Accessed: 10/04/2022.

**U.S. Government Accountability Office.** 2021. "Prescription Drugs: Medicare Spending on Drugs with Direct-to-Consumer Advertising." *https://www.gao.gov/products/gao-21-380*, Accessed: 10/04/2022.

**Wang, Junling, Ilene H. Zuckerman, Nancy A. Miller, Fadia T. Shaya, Jason M. Noel, and C. Daniel Mullins.** 2007. "Utilizing New Prescription Drugs: Disparities among Non-Hispanic Whites, Non-Hispanic Blacks, and Hispanic Whites." *Health Services Research*, 42(4): 1499–1519.

**Welsh, John, Yuan Lu, Sanket S. Dhruva, Behnood Bikdeli, Nihar R. Desai, Liliya Benchetrit, Chloe O. Zimmerman, Lin Mu, Joseph S. Ross, and Harlan M. Krumholz.** 2018. "Age of Data at the Time of Publication of Contemporary Clinical Trials." *JAMA Network Open*, 1(4): e181065.

**Wexler, Deborah.** 2022. "Initial Management of Hyperglycemia in Adults with Type 2 Diabetes Mellitus." In *UpToDate*. , ed. Theodore Post. Wolters Kluwer.

**Yahoo!Finance Historical Data.** 2022. "Moderna Therapeutics (MRNA)." *https://finance.yahoo.com/quote/MRNA/history?p=MRNA*.

**Zarin, Deborah A, Steven N Goodman, and Jonathan Kimmelman.** 2019. "Harms from Uninformative Clinical Trials." *JAMA*, 322(9): 813–814.

# Appendix

## Table of Contents


**A Supplementary Results** **A.2**

A.1 Conditioning on Patient Characteristics . . . . . A.2

A.2 Interaction Between Representation and Efficacy . . . . . A.3

**B Appendix Figures** **A.4**

**C Appendix Tables** **A.22**

**D Selected Quotes** **A.40**

**E Survey Instrument** **A.45**

E.1 Physician Survey Panel . . . . . A.45

E.2 Patient Survey Panel . . . . . A.45

E.3 Survey Materials . . . . . A.46

**F Model Appendix** **A.52**

F.1 Model Details . . . . . A.52

F.2 Further Results . . . . . A.52

F.3 Proofs . . . . . A.52

F.4 The Value and Cost of Recruiting Strategies to the Firm . . . . . A.53

F.5 Numerical Examples . . . . . A.54

**G Institutional Details** **A.55**

G.1 Policy Efforts to Improve Representation in U.S. Government Sponsored Clinical Trials Research . . . . . A.55

G.2 Background on ClinicalTrials.gov . . . . . A.56

**H Data** **A.59**

H.1 Clinical Trials . . . . . A.59

H.2 Prescribing Rates . . . . . A.60

H.3 Other Data Sources . . . . . A.62

# A Supplementary Results

## A.1 Conditioning on Patient Characteristics

We document, using survey data from Research!America, that Black patients are less likely to hear about clinical trials, less likely to enroll in trials if recommended by their doctor, and cite (a lack of) trust as one of the reasons for their hesitancy in enrolling (Table I). Overall, this evidence suggests that there is racial disparity in access to and perceived benefits from trials.

A standard question is whether these gaps persist when conditioning on patient characteristics. Black patients statistically differ in characteristics like income, education, and insurance status, and these characteristics also influence participation and perception of clinical trials. Without controlling for such characteristics, the gaps presented could suffer from omitted variable bias.

However, recent work in economics and law highlights that conditioning on patient characteristics may also result in "included variable bias" (Ayres 2010). The key distinction is that an agent's practices, which seem neutral when conditioning on patient characteristics, can result in "substantial adverse impact on a protected group." Even if there is no "intentional discrimination," such practices are discriminatory unless they are essential for the task at hand (Ayres 2010). For example, recruiting from academic medical centers and not safety net hospitals might be taken as a neutral business practice (conditional on hospital access) but if minority and immigrant communities use the safety net system more often this could have a disparate impact.

On the supply side of clinical trials, the question of which characteristics to condition on is subtle. For example, education may not directly affect the ability of a patient to participate in a trial, suggesting that we should *not* control for education. On the other hand, education might affect the ability to provide informed consent and therefore be an appropriate control.[66] Similar arguments follow for prescription behavior — on the one hand, insurance and income often determine access to new medications.[67] On the other hand, the presence of systemic racial barriers in access to employment and insurance, including importantly a lack of universal health care and other safety net systems in the US, might suggest that income and insurance are mechanisms of discrimination.[68]

Empirically, we find that conditional or unconditional gaps are quite similar for clinical trial participation and new drug prescription rates. In Appendix Table C1, we show that the gaps in access to trials and beliefs on benefits from trials are unchanged when we control for income, education, and political affiliation. Regardless, the notion that conditional gaps are the "right" measure of disparity requires careful evaluation and should not be asserted without considering the specific context.

[66] Though the response to such a rationale might very well be that consent forms could be accessible to people with different literacy levels.

[67] Indeed, the two are linked in the U.S. through employer-sponsored health insurance.

[68] According to Alesina, Glaeser and Sacerdote (2001): "Within the United States, race is the single most important predictor of support for welfare. America's troubled race relations are clearly a major reason for the absence of an American welfare state."

## A.2 Interaction Between Representation and Efficacy

The effect of representation on prescription intent may differ based on the efficacy of the drug. Our model suggests that representation matters the least when a drug has very low or very high efficacy. Intuitively, if a drug is ineffective for all patients in a trial, then a physician will not prescribe it to Black patients, independent of their decision-making with regard to representation. Similarly, if a drug is a drastic improvement over existing treatments, then a physician will be willing to prescribe to all patients, even those belonging to underrepresented groups. For more intermediate ranges of efficacy, we would expect representation to meaningfully impact a physician's prescription intent.

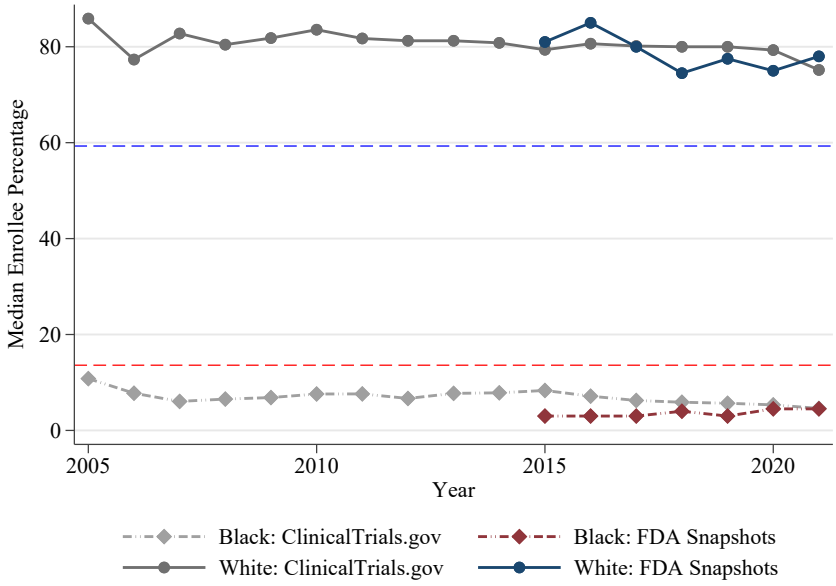
In our experiment, we choose the domain of efficacy values to reflect typical values of FDA-approved oral antiglycemics. The efficacy range of these drugs — 0.5 to 2 percentage point reductions in A1c — is narrow. Although this choice allowed us to study our central question in a way that simulates real-world physician decisions, it did limit our ability to detect potential differences in the effect of representation on prescription intent across different values of efficacy.

In Column (2) of Table C14, we observe that the coefficient on the interaction between efficacy and representation is not economically meaningful nor is it statistically significant. This confirms that the effect of representation on prescription intent is essentially constant across the narrow range of efficacy we present in the experiment.

We also note that alternative models would predict that the effect of representation on prescription intent is identical across all domains of efficacy. For example, in models of "warm-glow giving" (Andreoni 1990), physicians (agents) directly derive utility from prescribing drugs tested in representative trials (donating to charity) independent of the reported efficacy (amount of donation). Investigating whether the effect of representation on prescription intent varies for a larger domain of efficacy values is an avenue for future research.

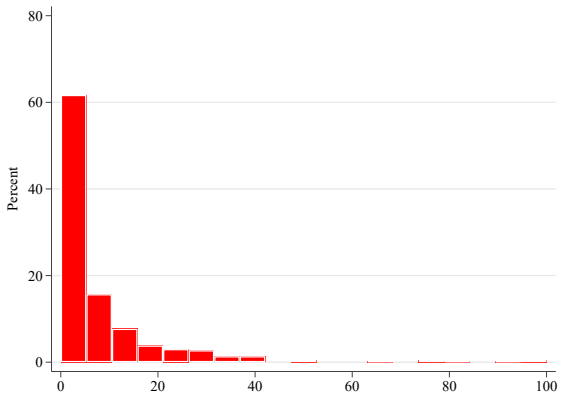
# B Appendix Figures

Appendix Figure B1: Clinical Trials Participation in ClinicalTrials.gov and FDA Drug Trials Snapshots

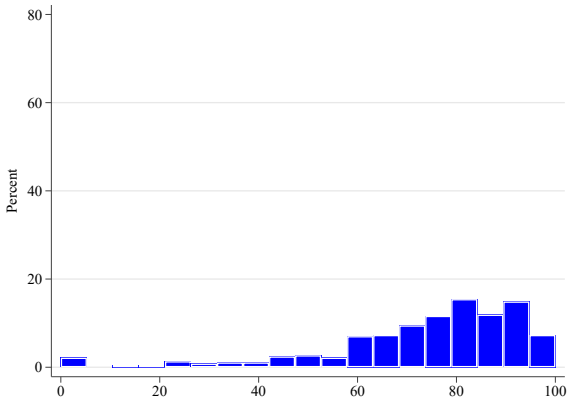



*Notes:* Figure plots the median enrollee shares by race in clinical trials, using data drawn from two databases: FDA Drug Trials Snapshots and ClinicalTrials.gov. FDA Drug Trials Snapshots includes data on all pivotal trials for drugs approved between 2015 and 2021. ClinicalTrials.gov includes data on all trials for completed trials that report Black and/or White enrollment rates, with a primary completion date between 2005 and 2021. Dashed lines plot the population shares by race in the U.S. population as reported in the 2020 U.S. Census (Black population share is 13.6 percent and non-Hispanic White population share is 59.3 percent; U.S. Census Bureau 2021). See Appendix H.1.1 and H.1.2 for details on data.

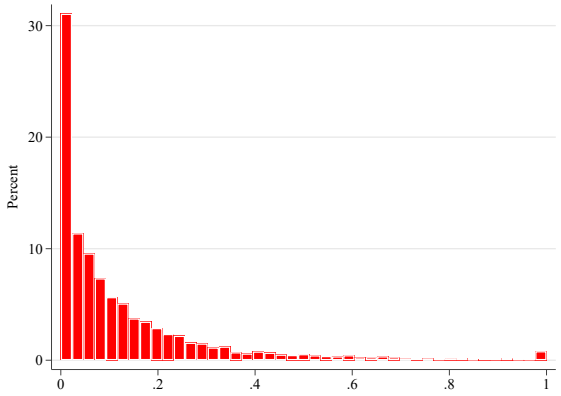
Appendix Figure B2: Trial Participation By Race



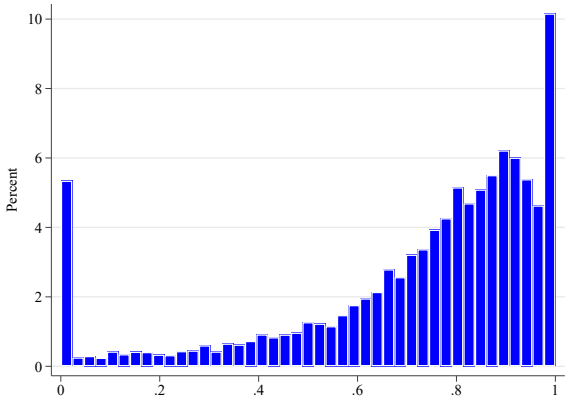

(a) FDA Drug Trials Snapshots, Black Patients




(b) FDA Drug Trials Snapshots, White Patients



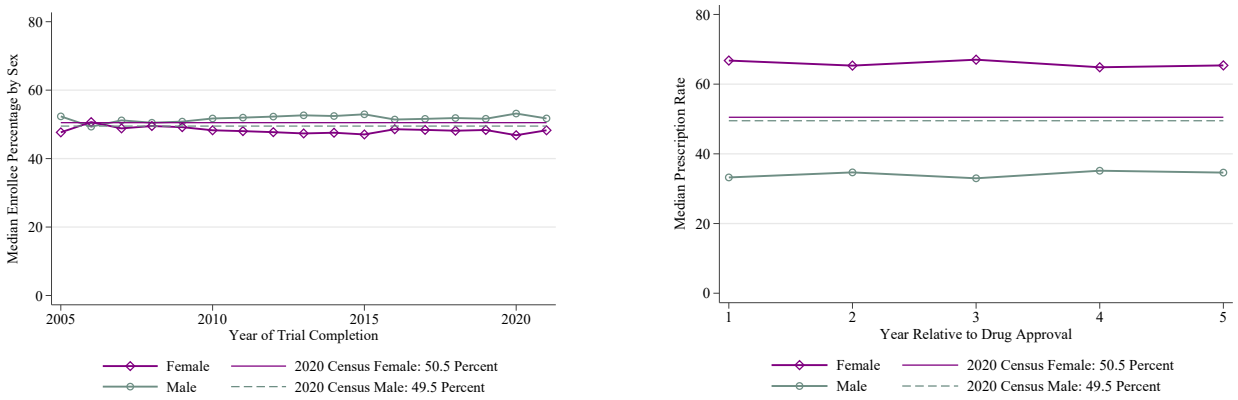

(c) ClinicalTrials.gov, Black Patients




(d) ClinicalTrials.gov, White Patients

*Notes:* Figure plots the racial composition of clinical trials separately for Black and White participants. Panels (a) and (b) use data drawn from the FDA Drug Trials Snapshots database. Panels (c) and (d) use data from ClinicalTrials.gov for completed trials that report Black and/or White enrollment rates. See Appendix H.1.1 and H.1.2 for details on data.

Appendix Figure B3: Development and Distribution of New Drugs by Sex

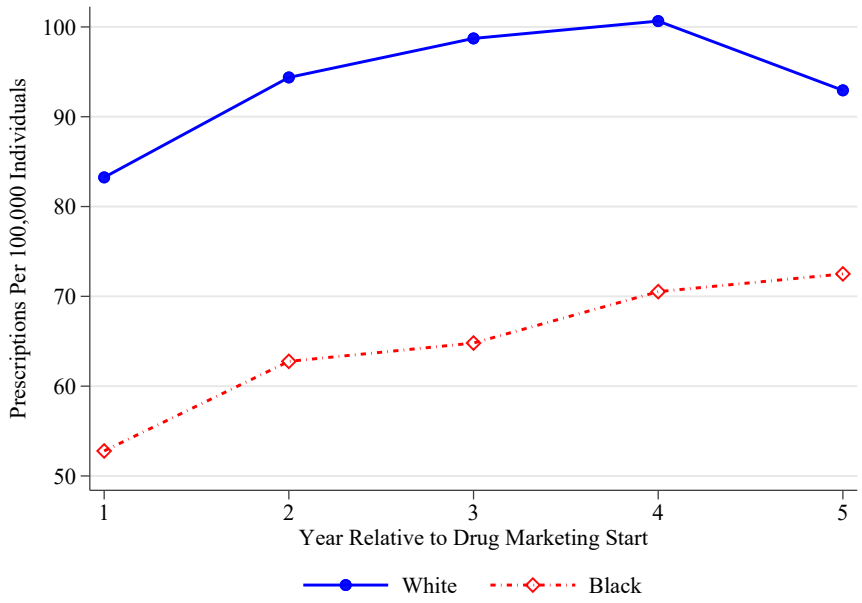



(a) Clinical Trials Participation

(b) Prescriptions of New Drugs

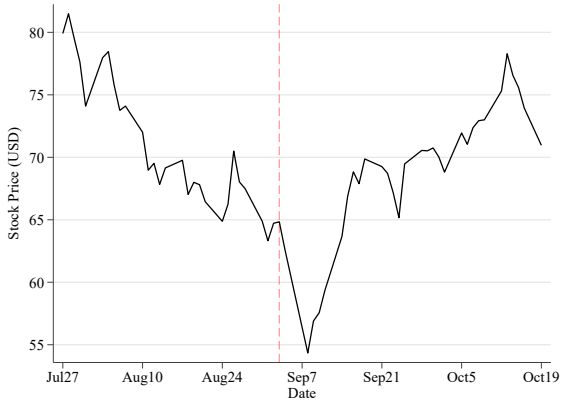
*Notes:* Figure replicates Figure I using data on sex instead of race. Panel (a) plots the median enrollee shares by sex using data drawn from ClinicalTrials.gov for completed trials that report data on sex and a primary completion date between 2005 and 2020. Dashed lines plot population shares by sex in the U.S. in 2020. Panel (b) plots the median prescription rate of a new drug by sex in each year relative to its approval. Data are drawn from the Medical Expenditure Panel Survey. Dashed lines in both panels plot population shares by sex in the U.S. as reported in the 2020 Census (Female population share is 50.5 percent, and Male population share is 49.5 percent; U.S. Census Bureau 2021). See Appendix H.1.1, H.2, and H.3.3 for details on data.

Appendix Figure B4: Prescribing Rates per Population

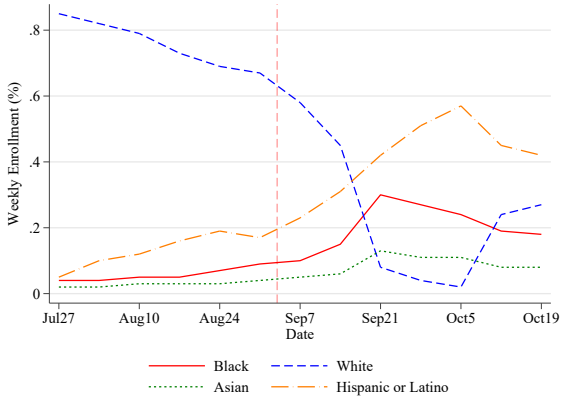



*Notes:* Figure plots the average number of prescriptions in each year relative to new drug marketing start date per 1000 individuals. The average number of prescriptions is plotted separately for Black and White individuals. Data are drawn from the Medical Expenditure Panel Survey. See Appendix H.2 for details on data.

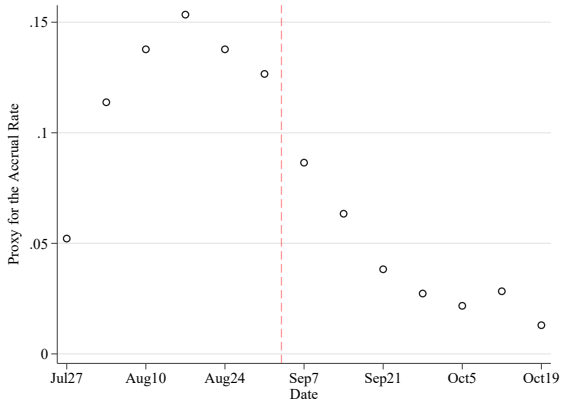
Appendix Figure B5: Moderna Stock Price and Trial Enrollment




(a) Moderna Stock Price



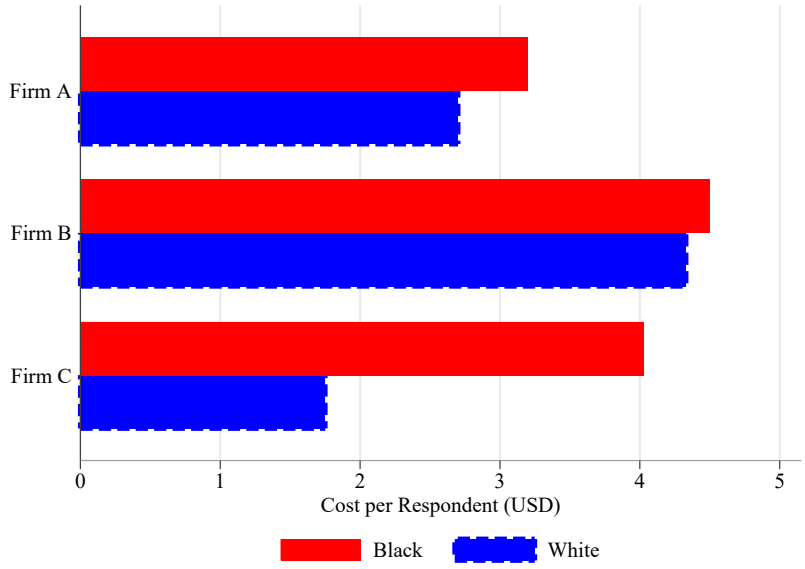

(b) Trial Enrollment by Race and Ethnicity




(c) Proxy for the Accrual Rate

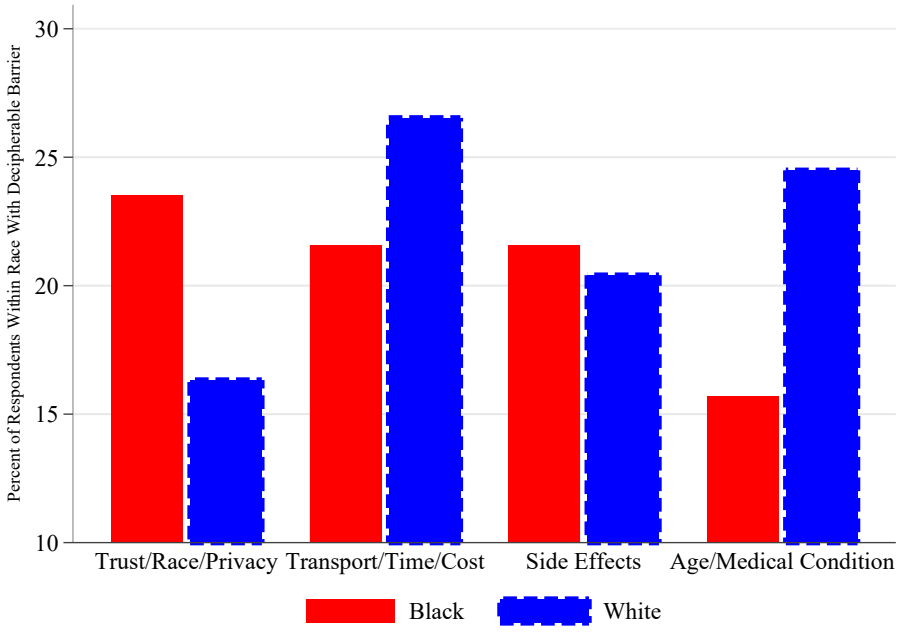
*Notes:* Panel (a) plots the Moderna Inc. stock price from July 27, 2020 through October 19, 2020. Data are from Yahoo!Finance Historical Stock Records. Panel (b) plots the share of new Moderna Covid-19 trial participants by race. Panel (c) plots a proxy for the accrual rate — the number of participants enrolled in the trial that week divided by the total trial enrollment, which was pre-specified by the company at 30,000 (National Institutes of Health 2020). Data for Panels (b) and (c) are from Moderna presentations and executive announcements. In all panels, the vertical line at September 3, 2020 marks the date of Moderna's public announcement of slowing down trial enrollment to ensure minority representation (Tirrell and Miller 2020). See Appendix H.3.6 and H.3.7 for details on data.

Appendix Figure B6: Recruitment Costs by Race across Firms



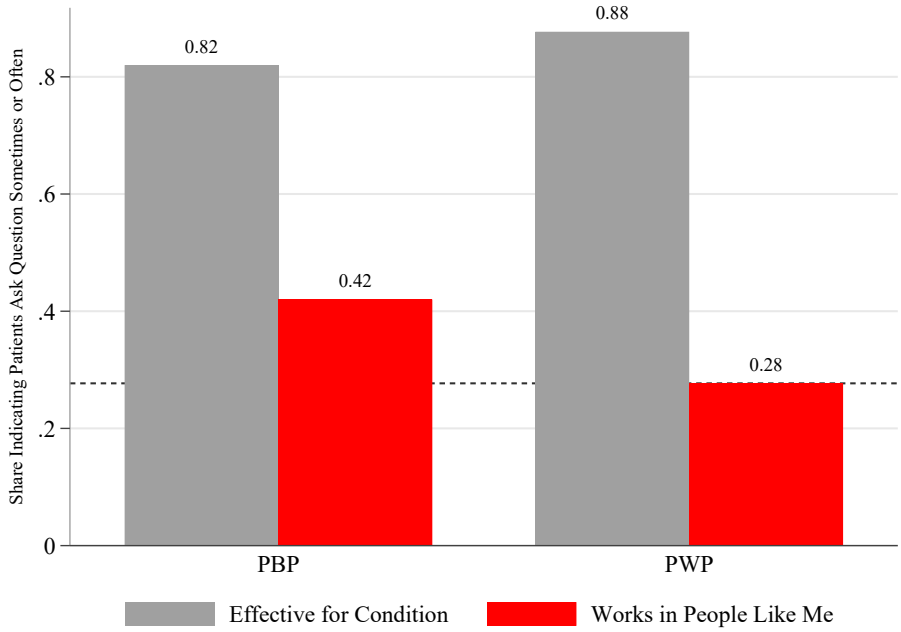

*Notes:* Figure plots the estimated per-respondent cost of recruiting survey participants on three online platforms. Quotes are provided by race for samples of 400 respondents (Firm A) and 600 respondents (Firms B and C). Data are from estimates solicited by the authors between January and June of 2022 from large marketing research firms, which have been used to recruit participants for online experiments in economics; firm names have been anonymized.

Appendix Figure B7: Leading Barriers to Clinical Trial Participation



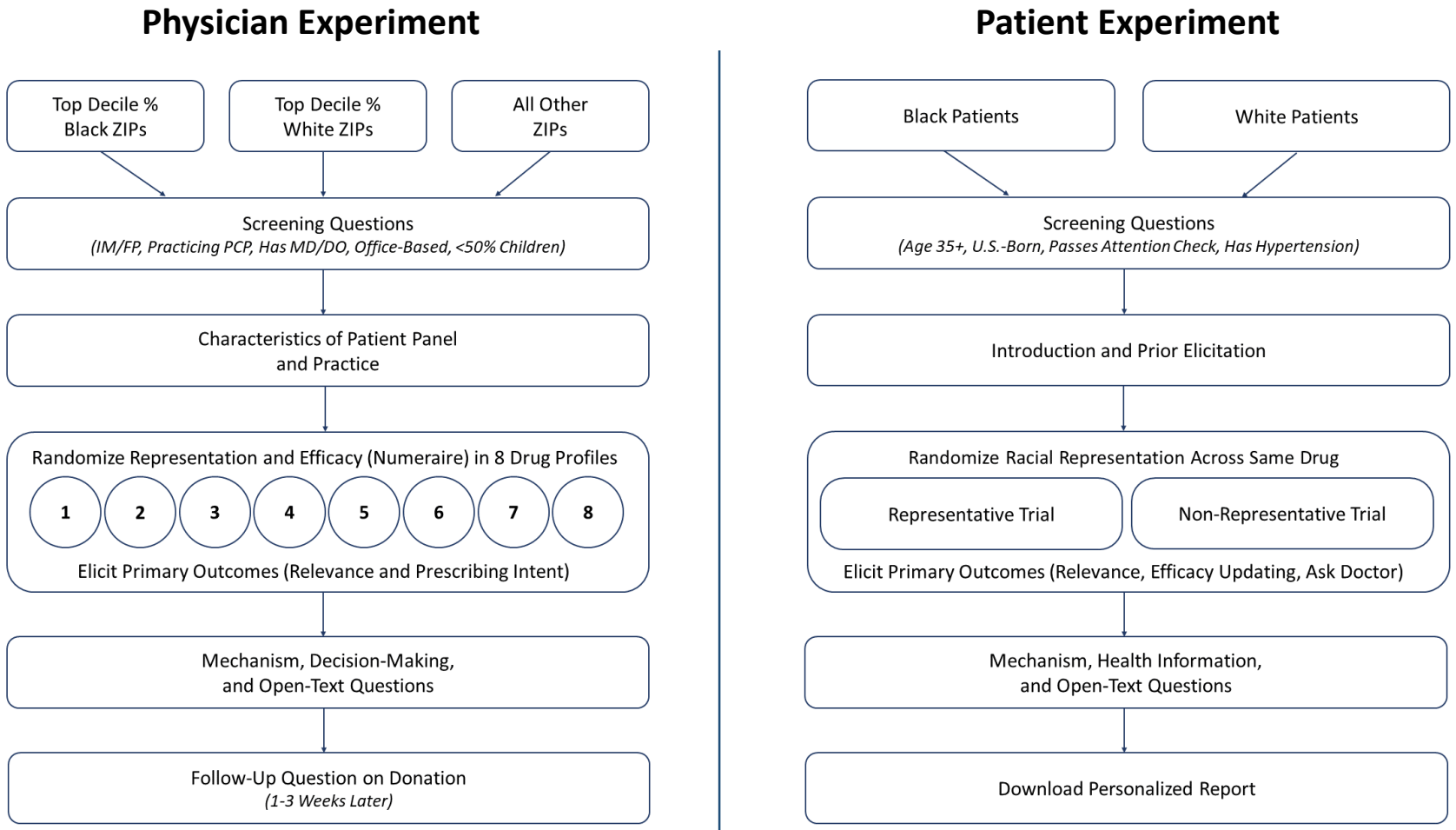

*Notes:* Figure presents the most commonly cited barriers to participating in clinical trials among patient respondents reporting they faced a decipherable barrier to participation, by respondent race. Open-text responses were independently coded by three coders (two research assistants and one graduate student) as corresponding to one of eight categories of barriers: (1) trust/race/privacy; (2) transport/time/cost; (3) side effects; (4) age/medical condition; (5) lack information; (6) not interested; (7) none/no barriers; and (8) indecipherable/unsure. In the event of disagreement between coders, the code selected by the majority was used; in the rare (N=4) event all coders disagreed, one of the codes was selected at random. In total, 36.0 percent of Black respondents and 36.8 percent of White respondents reported facing a decipherable barrier to trial participation; as visualized in the figure, the older age profile of White respondents corresponds to greater age-related barriers. Data are from the Patient Survey Experiment.

Appendix Figure B8: Science Research Perceived Benefit by Demographic Group



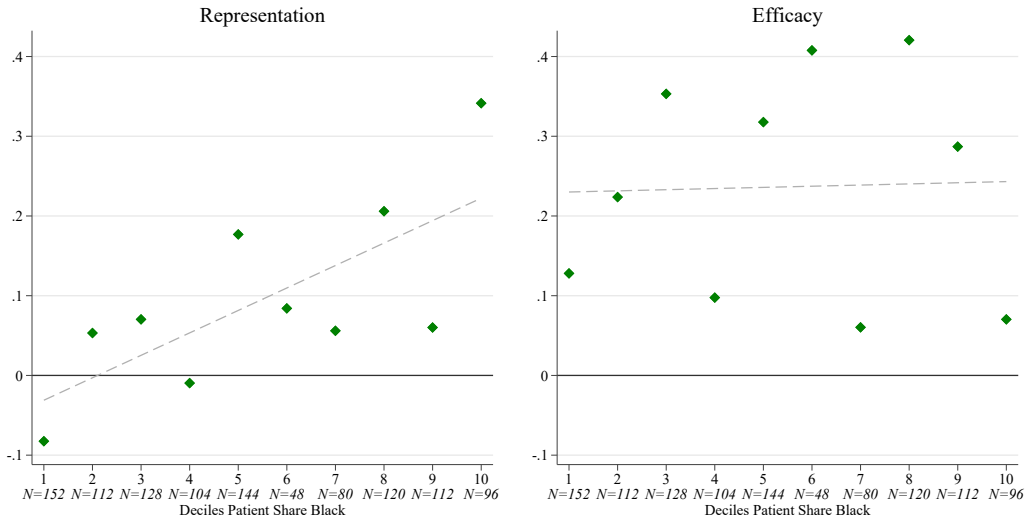

*Notes:* Figure plots the share of physicians who indicate that they have been asked the following questions by patients: "Is the drug effective for my condition?" and "How do I know the drug will work in people like me?" PBP (Physicians treating Black patients) denotes physicians who report above the median percent Black patients in their patient panel. PWP (Physicians treating White patients) is defined similarly with respect to White patients. Data are from the Physician Survey Experiment.

Appendix Figure B9: Physician and Patient Experiment Flow



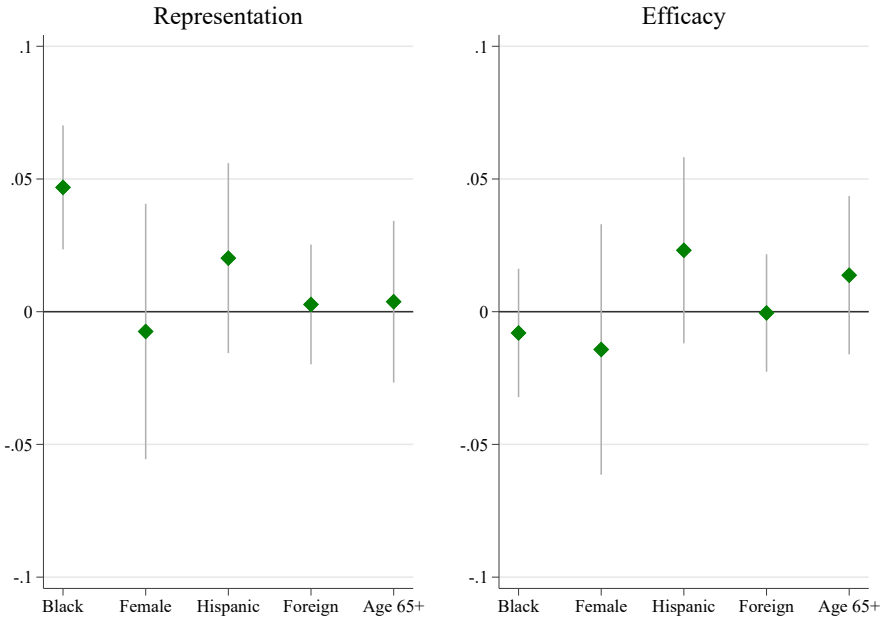

*Notes:* Overview of the Physician and Patient Survey Experiments.

Appendix Figure B10: Prescribing Willingness by Deciles of Patient Percent Black



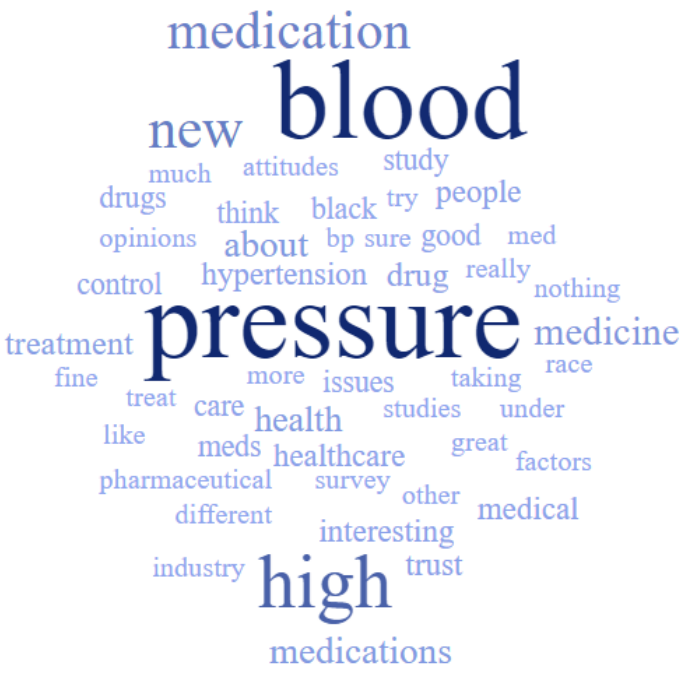

*Notes:* Figure plots the relationship between trial *Efficacy* and *Prescribing Intention* (on a 0-10 scale) by patient composition and percent Black of trial subjects in the profiles shown to physicians.

Appendix Figure B11: Association Between Physician Coefficients and Patient Characteristics



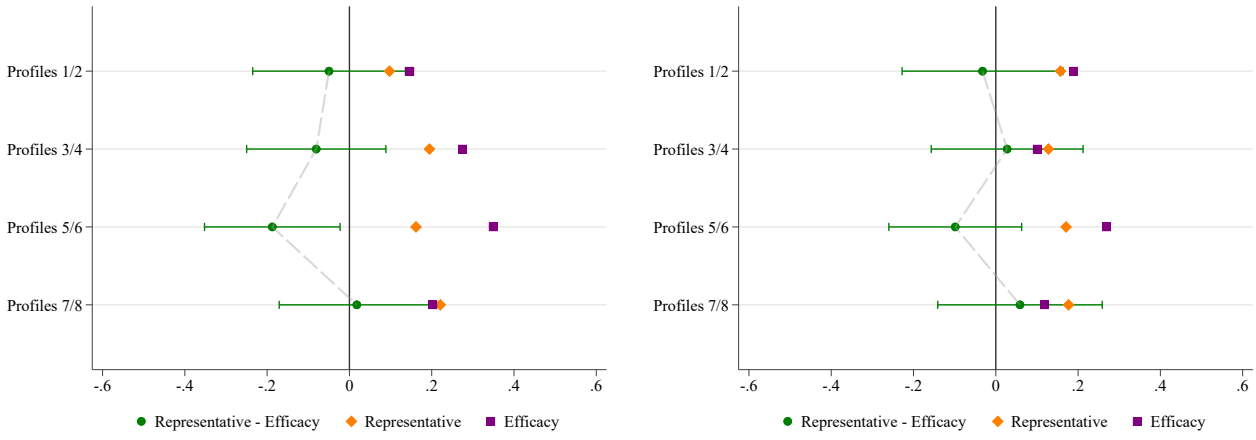

*Notes:* Figure plots coefficient estimates from regressions of physician-specific coefficients for representative treatment or efficacy treatment on patient panel characteristics (expressed as the percentage of patients with a given demographic characteristic multiplied by 10). 95 percent confidence intervals using robust standard errors are drawn. Data are from the Physician Survey Experiment.

Appendix Figure B12: Patient Respondents: Open-Text Responses




*Notes*: Word cloud of open-text responses of patient respondents' answers to the question at the end of the survey: "What do you think this study is about?" Data are drawn from the Patient Survey Experiment.

Appendix Figure B13: Physician Experimental Results by Profile Order

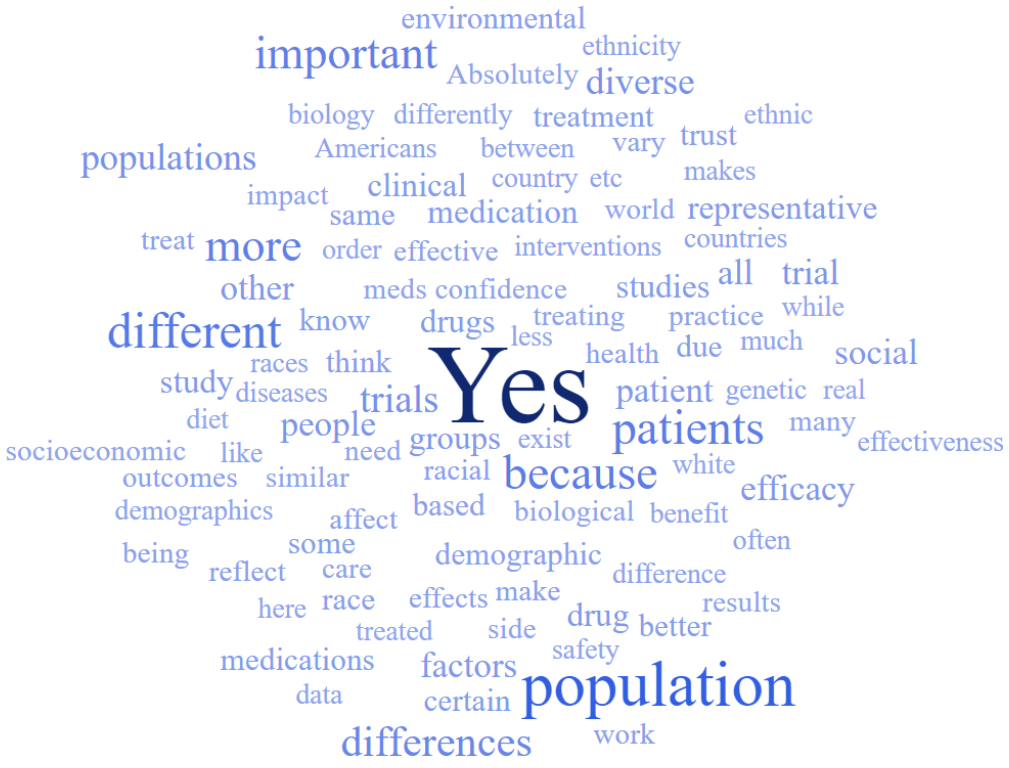



(a) Prescribing

(b) Relevance

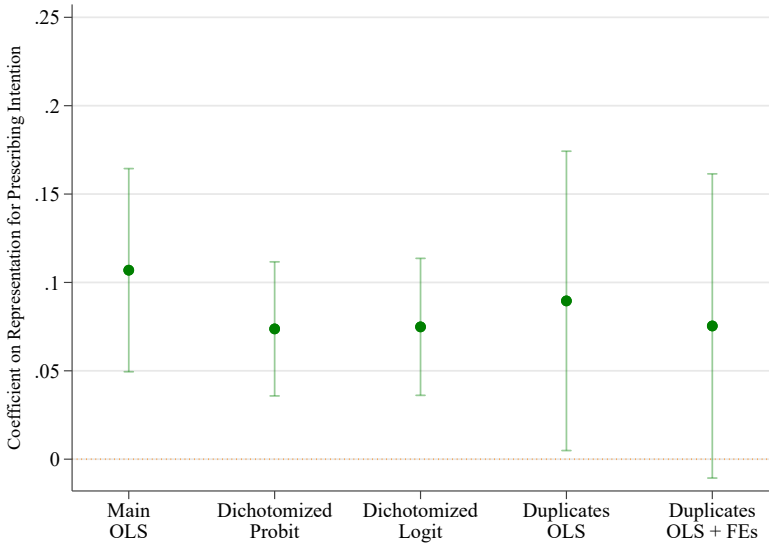
*Notes:* Figure plots the difference of coefficients on representative treatment and efficacy treatment as well as their point estimates for the outcomes of *Relevance* and *Prescribing Intention*, grouped by profile orders. *Representative* refers to the randomized percent Black in the trial unless otherwise indicated. *Efficacy* refers to the randomized percentage point drop in A1c. *Prescribing Intention*, *Representative* and *Efficacy* are standardized to a mean of 0 and a standard deviation of 1. Rx Mechanism fixed effects are included. Robust standard errors are clustered at the physician level. 95 percent confidence intervals for difference estimates are displayed.

Appendix Figure B14: Physician Respondents: Open-Text Responses



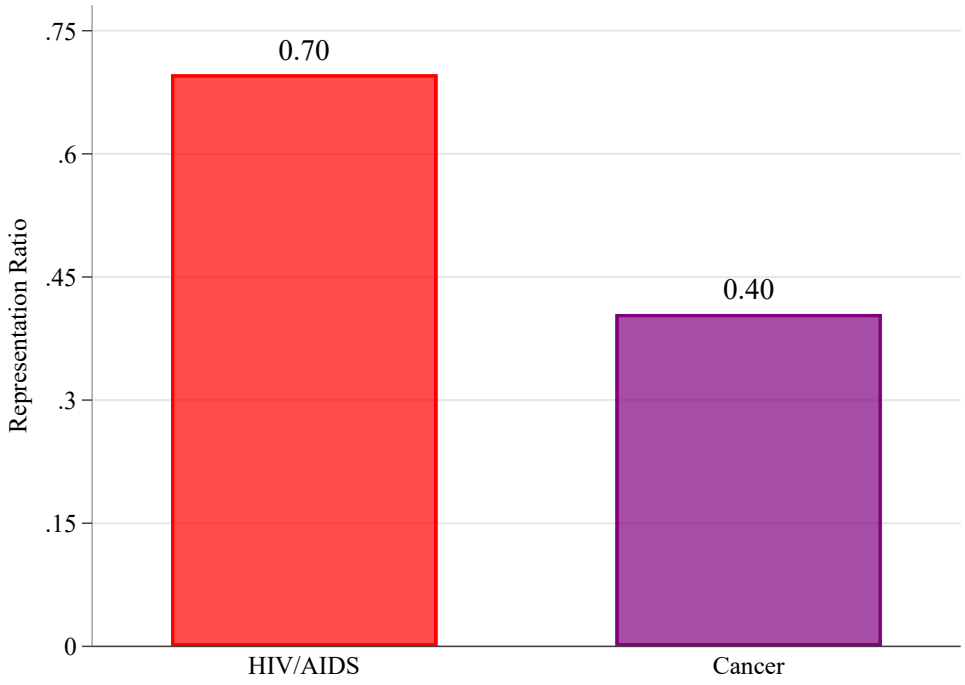

*Notes:* Word Cloud of open-text responses of physicians' answers to the question: "Is it important for clinical trials to be representative of the U.S. population? Why or why not?" Data are from the Physician Survey Experiment.

Appendix Figure B15: Robustness Across Samples and to Alternate Specifications: Prescribing Intention



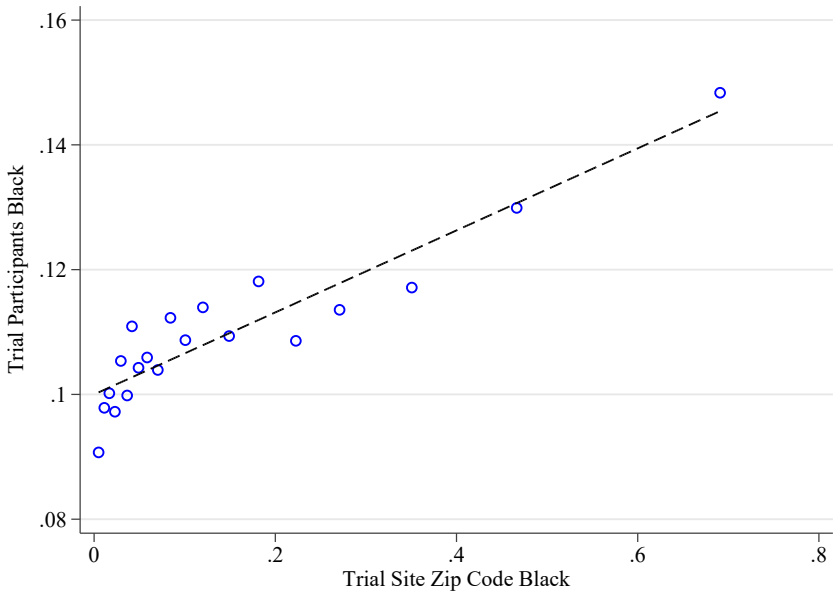

*Notes:* Figure plots coefficient estimates from a regression of prescribing intention on representation. *Representation* is standardized for all specifications; *Prescribing Intention* is standardized to a mean of 0 and standard deviation of one for OLS specifications and dichotomized at the median value for probit and logit specifications. The "Main OLS" specification corresponds to our pre-specified main equation. The "Dichotomized Probit" specification corresponds to a probit equation regressing dichotomized prescribing intention on *Representation* and *Efficacy* with profile order fixed effects and drug mechanism fixed effects, using robust standard errors clustered at the physician level. The "Dichotomized Logit" specification is identical to the probit specification but logistic regression is utilized instead. In both the probit and logit cases, the margin effects are plotted. The "Duplicates OLS" specification plots estimates from a regression of *Prescribing Intention* on *Representation* with physician-by-efficacy fixed effects. Within our sample, we ensured that each physician saw two drug profiles with identical efficacy levels and different measures of representation; "Duplicates OLS" restricts consideration to this sample. The "Duplicates OLS + FEs" specification is identical to the "Duplicates OLS" but adds profile order fixed effects and drug mechanism fixed effects. 95 percent confidence intervals are shown; all coefficient estimates are statistically significant at the 95 percent level except "Duplicates OLS + FEs," which is significant at the 90 percent level. Data are from the Physician Survey Experiment.

Appendix Figure B16: Representation Relative to Disease Burden



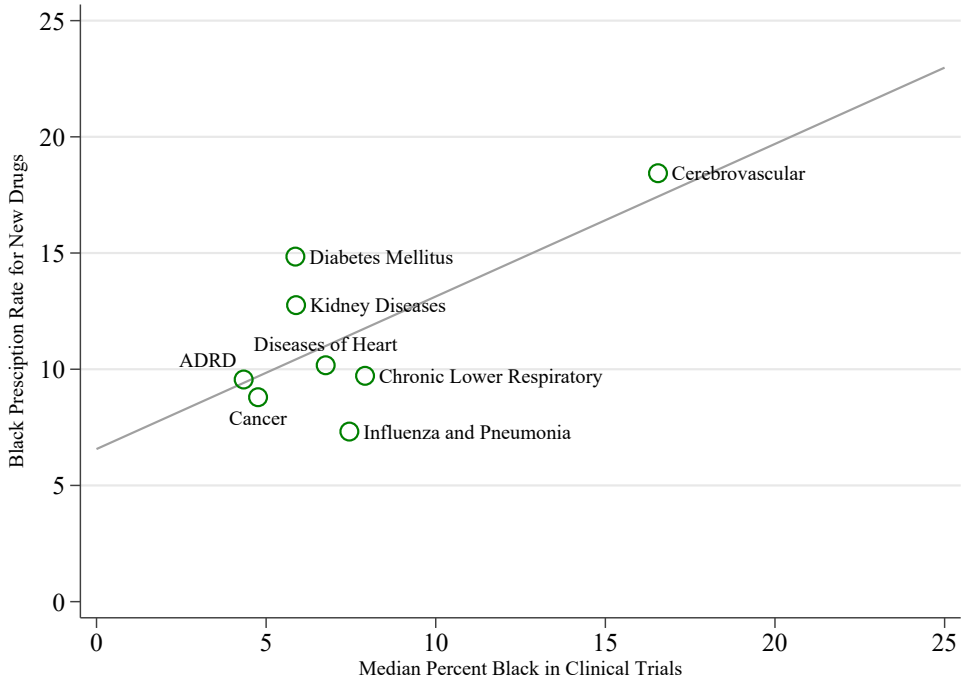

*Notes:* Figure plots the representation ratios of cancer and HIV/AIDS in Black patients relative to disease burden. Representation ratios are defined as the median percent Black in trials in a disease category divided by the Black disease burden (share of Black deaths among all deaths from the condition in the United States). Enrollment data are drawn from ClinicalTrials.gov for completed trials that report Black enrollment rates. Black HIV/AIDS disease burden was 40.2 percent in 2019 (Centers for Disease Control and Prevention 2021). Black cancer disease burden was 11.8 percent in 2019 (Heron 2021). See Appendix H.1.1 for details on data.

Appendix Figure B17: Racial Composition of Clinical Trials and Trial Site Zip Codes




*Notes:* Figure plots the binned average percent Black in clinical trials and the average percent Black of trial site zip codes. Enrollment data are drawn from ClinicalTrials.gov for completed trials that report Black enrollment rates and have trial sites in the United States. Data on trial site zip code demographics are drawn from the 2019 American Community Survey (ACS). See Appendix H.1.1 and H.3.3 for details on data.

Appendix Figure B18: Prescription Gaps and Trial Inclusivity (Excluding HIV-AIDS)




*Notes:* Figure plots the correlation between the prescription rate of new medications to Black Americans and the median percent Black in pivotal trials. We construct the prescription rate as the percentage of newly marketed drugs (on the market for five or fewer years) received by Black Americans in the ten leading causes of death (excluding unintentional injuries and suicide) in the United States (Heron 2021). The y-axis value of "Cancer" refers to outpatient cancer supportive therapies. Prescription data are drawn from the Medical Expenditure Panel Survey, and trial composition data are drawn from ClinicalTrials.gov. We include a line of best fit. See Appendix H.1.1 and H.2 for details on data construction.

# C Appendix Tables

Appendix Table C1: Polling Data on Clinical Trial Barriers and Beliefs

| | *Confidence in Research Institutions* | | *Heard of Clinical Trial* | | *Would Enroll if Doctor Recommends* | | *Trust Not Reason for Lack of Enrollment* | | *Science is Beneficial* | | *Would Get FDA-Approv. Vaccine* | | *Kling-Liebman-Katz (KLK) Index* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Black | -0.253** | -0.273** | -0.079*** | -0.076*** | -0.054** | -0.046* | -0.104*** | -0.090*** | -0.099** | -0.106** | -0.163 | -0.054 | -0.226*** | -0.215*** |
| | (0.115) | (0.123) | (0.020) | (0.022) | (0.022) | (0.023) | (0.029) | (0.032) | (0.045) | (0.046) | (0.119) | (0.128) | (0.022) | (0.024) |
| Constant | 3.082 | 3.329 | 0.875 | 0.909 | 0.837 | 0.871 | 0.536 | 0.528 | 0.383 | 0.595 | 3.069 | 3.436 | 1.152 | 1.218 |
| Covariates | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 940 | 927 | 2,843 | 2,757 | 2,658 | 2,584 | 2,031 | 1,948 | 971 | 955 | 922 | 907 | 2,968 | 2,874 |

*Notes:* Table reports OLS estimates from a regression of survey responses among Black and White individuals across a number of questions regarding science. Covariates include an indicator for whether the individual's income is above the median, an indicator for whether the individual has a college degree, and an indicator for whether the individual supports GOP. Data is drawn from a nationally representative survey conducted by Research!America in the years 2013, 2017, and 2021. See Appendix H.3.5 for details on data. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C2: Physician Survey Experiment Balance Table

| | *Mean of Values Over Trials* | | *Range of Values Over Trials* | |
|---|---|---|---|---|
| | Representation | Efficacy | Representation | Efficacy |
| | (1) | (2) | (3) | (4) |
| Physician Age | 0.017 | 0.007 | -0.094 | -0.002 |
| | (0.015) | (0.014) | (0.094) | (0.003) |
| Physician is Male | 0.123 | -0.344* | 0.912 | -0.009 |
| | (0.185) | (0.191) | (1.082) | (0.037) |
| Physician is White | -0.108 | 0.017 | 0.300 | -0.051 |
| | (0.201) | (0.206) | (1.111) | (0.041) |
| Physician Hours/Week | -0.009 | 0.005 | -0.039 | -0.003** |
| | (0.007) | (0.005) | (0.037) | (0.001) |
| Physician Years Practice (Grp) | -0.090 | -0.060 | -0.320 | 0.011 |
| | (0.108) | (0.103) | (0.650) | (0.023) |
| Physician Holds MD | 0.164 | -0.088 | -0.627 | 0.013 |
| | (0.271) | (0.261) | (1.445) | (0.059) |
| Patient Percent Black | 0.007 | 0.005 | 0.052 | -0.001 |
| | (0.006) | (0.006) | (0.038) | (0.002) |
| Patient Percent White | 0.009 | -0.001 | 0.060* | >-0.001 |
| | (0.006) | (0.006) | (0.034) | (0.002) |
| Patient Percent Hispanic | 0.008 | 0.002 | 0.062 | -0.001 |
| | (0.007) | (0.007) | (0.041) | (0.002) |
| Altruism (0-10) | 0.014 | 0.001 | 0.482 | -0.021 |
| | (0.069) | (0.081) | (0.402) | (0.015) |
| Risk Preference (0-10) | 0.037 | 0.071* | -0.357 | 0.003 |
| | (0.047) | (0.042) | (0.283) | (0.009) |
| Time Preference (0-10) | 0.071 | -0.021 | 0.452 | 0.012 |
| | (0.080) | (0.069) | (0.443) | (0.015) |
| Constant | -1.960* | -0.356 | 22.475*** | 1.423*** |
| | (1.043) | (0.883) | (5.430) | (0.238) |
| Observations | 137 | 137 | 137 | 137 |

*Notes:* Table displays results from separate regressions of the mean and range of representation and efficacy values randomly assigned to physicians on a host of physician and physicians' patient panel characteristics. For Physician Years Practice, respondents selected an interval from a multiple choice list; "Grp" refers to the group (selected interval) chosen by the respondent, instead of the numerical value indicated. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

## Appendix Table C3: Characteristics of Physicians and Their Communities: Comparison Between Participants and Physicians Nationwide by Strata

| | *Top Decile Share Black ZIPs* | | *Bottom Decile Share Black ZIPs* | | *All Other ZIPs* | | *Differences* | | |
|---|---|---|---|---|---|---|---|---|---|
| | AMA Physicians | Survey Respondents | AMA Physicians | Survey Respondents | AMA Physicians | Survey Respondents | Top Decile Black ZIPs | Bottom Decile Black ZIPs | All Other ZIPs |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Phys: Male | 0.548 | 0.559 | 0.618 | 0.543 | 0.569 | 0.558 | -0.011 | 0.075 | 0.010 |
| | (0.498) | (0.501) | (0.486) | (0.505) | (0.495) | (0.502) | (0.065) | (0.084) | (0.076) |
| Phys: Age | 44.587 | 49.254 | 48.388 | 48.543 | 46.470 | 50.349 | -4.667*** | -0.155 | -3.879** |
| | (10.948) | (10.405) | (10.464) | (10.239) | (10.488) | (10.433) | (1.346) | (1.709) | (1.573) |
| Phys: Yrs Since Deg | 16.827 | 16.275 | 19.622 | 15.310 | 18.386 | 17.711 | 0.552 | 4.312** | 0.676 |
| | (10.953) | (10.398) | (10.424) | (9.332) | (10.597) | (9.016) | (1.444) | (1.706) | (1.444) |
| Phys: Med School Rank | 99.205 | 67.745 | 84.494 | 79.448 | 90.693 | 85.763 | 31.460*** | 5.045 | 4.930 |
| | (37.596) | (46.052) | (41.779) | (45.826) | (40.724) | (43.509) | (6.392) | (8.374) | (6.965) |
| ZIP: South | 0.462 | 0.441 | 0.124 | 0.057 | 0.323 | 0.186 | 0.021 | 0.067* | 0.137** |
| | (0.499) | (0.501) | (0.329) | (0.236) | (0.468) | (0.394) | (0.065) | (0.039) | (0.059) |
| ZIP: Poverty Rate | 26.635 | 25.688 | 11.678 | 9.063 | 13.699 | 13.023 | 0.947 | 2.615*** | 0.676 |
| | (11.123) | (10.178) | (9.384) | (4.970) | (9.398) | (12.884) | (1.317) | (0.834) | (1.942) |
| ZIP: Black | 0.537 | 0.550 | 0.002 | 0.002 | 0.090 | 0.089 | -0.014 | 0.000 | 0.001 |
| | (0.207) | (0.211) | (0.002) | (0.002) | (0.094) | (0.093) | (0.027) | (0.000) | (0.014) |
| ZIP: Hispanic | 0.203 | 0.185 | 0.109 | 0.045 | 0.168 | 0.119 | 0.018 | 0.064*** | 0.049** |
| | (0.214) | (0.194) | (0.209) | (0.047) | (0.193) | (0.129) | (0.025) | (0.008) | (0.019) |
| ZIP: Asian | 0.041 | 0.047 | 0.016 | 0.018 | 0.077 | 0.081 | -0.006 | -0.002 | -0.004 |
| | (0.059) | (0.066) | (0.031) | (0.027) | (0.100) | (0.066) | (0.009) | (0.005) | (0.010) |
| ZIP: Age 18 and Under | 0.231 | 0.230 | 0.210 | 0.218 | 0.202 | 0.193 | 0.001 | -0.008 | 0.009 |
| | (0.053) | (0.044) | (0.059) | (0.047) | (0.060) | (0.059) | (0.006) | (0.008) | (0.009) |
| ZIP: Age 65 and Over | 0.127 | 0.130 | 0.208 | 0.207 | 0.158 | 0.161 | -0.003 | 0.002 | -0.003 |
| | (0.038) | (0.033) | (0.084) | (0.048) | (0.063) | (0.063) | (0.004) | (0.008) | (0.009) |
| ZIP: Insured | 0.885 | 0.888 | 0.933 | 0.952 | 0.926 | 0.945 | -0.003 | -0.019*** | -0.020*** |
| | (0.065) | (0.065) | (0.051) | (0.035) | (0.050) | (0.042) | (0.008) | (0.006) | (0.006) |
| Observations | 16,651 | 59 | 9,376 | 35 | 143,623 | 43 | 16,710 | 9,411 | 143,666 |

*Notes:* Table means and differences between physicians that participated in the survey experiment and all those coded as internal medicine or family practice in the U.S. who practice in similar zip codes. Data are drawn from the Physician Survey Experiment, U.S. News and World Report, the 2014 version of the AMA Masterfile Directory, and the 2019 American Community Survey. See Appendix H.3.1, H.3.2, and H.3.3 for details on data. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C4: Characteristics of Physicians Responding to Follow-Up Question

| | All Physicians (1) | Responded to Follow-Up (2) | Difference Between Groups (3) |
|---|---|---|---|
| Physician is Black | 0.088 | 0.085 | 0.002 |
| | (0.284) | (0.281) | (0.039) |
| Physician is White | 0.606 | 0.683 | -0.077*** |
| | (0.490) | (0.468) | (0.067) |
| Physician is Male | 0.555 | 0.585 | -0.031 |
| | (0.499) | (0.496) | (0.069) |
| Physician Age | 49.416 | 50.341 | -0.925* |
| | (10.319) | (9.918) | (1.420) |
| Physician is Republican | 0.190 | 0.159 | 0.031* |
| | (0.394) | (0.367) | (0.054) |
| Physician Hours/Week | 32.978 | 32.768 | 0.210 |
| | (13.740) | (13.159) | (1.889) |
| Physician Years Practice (Mdpt) | 16.460 | 17.293 | -0.833** |
| | (8.398) | (8.487) | (1.177) |
| MD Patients/Week (Mdpt) | 64.164 | 65.098 | -0.933 |
| | (30.941) | (30.872) | (4.316) |
| Patient Percent Black | 25.388 | 26.024 | -0.637 |
| | (23.131) | (23.792) | (3.264) |
| Patient Percent Female | 53.664 | 53.659 | 0.006 |
| | (11.858) | (11.708) | (1.648) |
| Patient Percent Children | 7.803 | 7.902 | -0.100 |
| | (8.058) | (7.780) | (1.111) |
| Patient Percent 65+ | 41.584 | 41.061 | 0.523 |
| | (18.727) | (16.604) | (2.508) |
| Patient Percent Foreign (Mdpt) | 27.591 | 26.037 | 1.555 |
| | (25.308) | (25.236) | (3.530) |
| Top Decile Black ZIP | 0.431 | 0.415 | 0.016 |
| | (0.497) | (0.496) | (0.069) |
| Bottom Decile Black ZIP | 0.255 | 0.280 | -0.025 |
| | (0.438) | (0.452) | (0.062) |
| Altruism (0-10) | 7.394 | 7.280 | 0.114 |
| | (1.447) | (1.468) | (0.203) |
| Risk Preference (0-10) | 5.730 | 5.683 | 0.047 |
| | (2.088) | (1.956) | (0.285) |
| Time Preference (0-10) | 7.854 | 7.927 | -0.073 |
| | (1.353) | (1.395) | (0.191) |
| Observations | 137 | 82 | 219 |

*Notes:* Table compares the baseline physician and patient panel characteristics of all physician respondents to those responding to the follow-up survey. For *Physician Years Practice (Mdpt)*, *Physician Patients/Week (Mdpt)*, and *Patient Percent Foreign (Mdpt)*, respondents were asked to select an interval from a list of options; we assigned each physician the *midpoint* of the interval selected when calculating summary statistics. Robust standard errors are used when comparing characteristics between the two groups. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C5: Comparison with 2019 MEPS on Variables in Both Surveys

| | *Non-Hispanic Black* | | | *Non-Hispanic White* | | |
|---|---|---|---|---|---|---|
| | MEPS | Survey | Difference | MEPS | Survey | Difference |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Male | 0.424 | 0.360 | 0.064 | 0.518 | 0.426 | 0.092** |
| | (0.494) | (0.482) | (0.044) | (0.500) | (0.496) | (0.043) |
| Age 45-64 | 0.498 | 0.482 | 0.016 | 0.411 | 0.382 | 0.029 |
| | (0.500) | (0.501) | (0.046) | (0.492) | (0.488) | (0.043) |
| Age 65+ | 0.386 | 0.295 | 0.091** | 0.508 | 0.478 | 0.030 |
| | (0.487) | (0.458) | (0.042) | (0.500) | (0.501) | (0.044) |
| BA or Higher | 0.194 | 0.331 | -0.136*** | 0.311 | 0.243 | 0.068* |
| | (0.396) | (0.472) | (0.042) | (0.463) | (0.430) | (0.038) |
| Under FPL | 0.385 | 0.374 | 0.011 | 0.216 | 0.279 | -0.063 |
| | (0.487) | (0.486) | (0.044) | (0.412) | (0.450) | (0.039) |
| Insured | 0.917 | 0.942 | -0.026 | 0.965 | 0.919 | 0.046* |
| | (0.277) | (0.234) | (0.022) | (0.184) | (0.274) | (0.024) |
| Observations | 1,153 | 139 | 1,292 | 4,146 | 136 | 4,282 |

*Notes:* Table compares the patient survey respondents, all of whom reported having hypertension, to individuals with hypertension in the 2019 Medical Expenditure Panel Survey. Survey weights are utilized. "Under FPL" refers to the household income of the respondent being under the 2021 federal poverty line for a four-person household (around $30k, see website of the Assistant Secretary for Planning and Evaluation). See Appendix H.2 for details on data. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C6: Patient Experimental Results — Robustness

**Panel A: Patients Demanding Personalized Report**

| | *Relevance* | | *Ask Doctor* | | *Loading on Signal* | |
|---|---|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Representative Treatment | 0.615** | 0.380 | 0.104 | 0.000 | 0.071 | -0.019 |
| | (0.258) | (0.253) | (0.113) | (0.126) | (0.106) | (0.059) |
| Observations | 63 | 52 | 63 | 52 | 63 | 52 |

**Panel B: Results Weighted Using MEPS Person Weights**

| | *Relevance* | | *Ask Doctor* | | *Loading on Signal* | |
|---|---|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Representative Treatment | 0.781*** | 0.166 | 0.042 | 0.008 | 0.132* | -0.076 |
| | (0.173) | (0.161) | (0.077) | (0.081) | (0.068) | (0.056) |
| Observations | 139 | 136 | 139 | 136 | 139 | 136 |

**Panel C: Inclusion of LASSO-Selected Controls**

| | *Relevance* | | *Ask Doctor* | | *Loading on Signal* | |
|---|---|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Representative Treatment | 0.781*** | 0.172 | 0.021 | 0.006 | 0.144** | -0.077 |
| | (0.164) | (0.158) | (0.077) | (0.079) | (0.066) | (0.056) |
| Observations | 139 | 136 | 139 | 136 | 139 | 136 |

*Notes: Relevance* is standardized to a mean of 0 and standard deviation of 1. Panel (a) reports estimates from Equation 3 restricting to patients demanding the personalized report. Panel (b) reports estimates weighting each observation based on the mean person weight in the 2019 Medical Expenditure Panel Survey by combinations of the following sociodemographic characteristics: race (Black or White), age (35–44, 45–64, or 65+), region (South vs. non-South), college degree, income (0-30k, 30-110k, 110k+), health insurance status (insured or uninsured), and usual place of care (has usual place or does not have usual place). Nine patient survey respondents have combinations of characteristics not represented in the MEPS, and are assigned person weight equal to the mean person weight in the MEPS. Results are robust to dropping these individuals (available upon request). Panel (c) reports estimates from double-selection LASSO linear regression. Potential controls included age, sex, education, and health variables among others. Robust standard errors clustered in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C7: Patient Survey Balance Table

| | All Respondents | Representative Arm | Non-Representative Arm | Difference |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Black | 0.505 | 0.518 | 0.493 | 0.025 |
| | (0.501) | (0.501) | (0.502) | (0.061) |
| Male | 0.393 | 0.388 | 0.397 | -0.009 |
| | (0.489) | (0.489) | (0.491) | (0.059) |
| Age Group | 5.876 | 5.914 | 5.838 | 0.075 |
| | (1.117) | (1.126) | (1.110) | (0.135) |
| BA or Higher | 0.287 | 0.281 | 0.294 | -0.014 |
| | (0.453) | (0.451) | (0.457) | (0.055) |
| Insured | 0.931 | 0.942 | 0.919 | 0.023 |
| | (0.254) | (0.234) | (0.274) | (0.031) |
| Takes BP Medication | 0.889 | 0.891 | 0.887 | 0.003 |
| | (0.315) | (0.313) | (0.318) | (0.038) |
| Past Nonadherence | 0.171 | 0.194 | 0.147 | 0.047 |
| | (0.377) | (0.397) | (0.355) | (0.045) |
| General Trust | 0.527 | 0.540 | 0.515 | 0.025 |
| | (0.500) | (0.500) | (0.502) | (0.060) |
| Pharma Trust | 1.636 | 1.669 | 1.603 | 0.066 |
| | (0.801) | (0.880) | (0.713) | (0.097) |
| Doctor Trust | 2.324 | 2.309 | 2.338 | -0.029 |
| | (0.689) | (0.700) | (0.680) | (0.083) |
| Public Health Trust | 1.945 | 1.971 | 1.919 | 0.052 |
| | (0.863) | (0.908) | (0.817) | (0.104) |
| Altruism | 6.793 | 6.748 | 6.838 | -0.090 |
| | (2.188) | (2.123) | (2.258) | (0.264) |
| Risk Preference | 5.422 | 5.273 | 5.574 | -0.300 |
| | (2.516) | (2.612) | (2.415) | (0.304) |
| Time Preference | 6.993 | 6.914 | 7.074 | -0.160 |
| | (1.985) | (2.094) | (1.872) | (0.240) |
| Heard of Tribenzor | 0.047 | 0.058 | 0.037 | 0.021 |
| | (0.213) | (0.234) | (0.189) | (0.026) |
| Prior on Efficacy | 5.782 | 5.928 | 5.632 | 0.296 |
| | (7.131) | (7.489) | (6.770) | (0.861) |
| Observations | 275 | 139 | 136 | 275 |

*Notes:* Table compares the baseline demographic, economic, and health characteristics of respondents receiving information on a representative trial (15 percent Black) to those receiving information on a non-representative trial (<1 percent Black).

Appendix Table C8: Patient Survey Experiment Attrition by Arm

| | Attrition Post-Randomization | |
|---|---|---|
| | (1) | (2) |
| Representative Treatment | 0.011 | |
| | (0.030) | |
| Black | | 0.036 |
| | | (0.030) |
| Sample Mean | 0.074 | 0.074 |
| Observations | 297 | 297 |

*Notes:* Table reports OLS estimates regressing an indicator for *Attrition Post-randomization* on indicators for *Representative Treatment* (*i.e.*, assignment to the 15 percent Black trial) and Black race. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C9: Physician Survey Experiment Attrition and Randomized Attributes

| | *Mean of Values Over Trials* | | *Range of Values Over Trials* | |
|---|---|---|---|---|
| | Representation | Efficacy | Representation | Efficacy |
| | (1) | (2) | (3) | (4) |
| Attrition | -0.121 | -0.099 | 0.911 | 0.014 |
| | (0.215) | (0.363) | (1.550) | (0.090) |
| Sample Mean | 0.000 | 0.000 | 25.782 | 1.159 |
| Observations | 145 | 145 | 145 | 145 |

*Notes:* Table displays results from separate regressions of the mean and range of *Representation* and *Efficacy* values randomly assigned to physicians on an indicator for respondent attrition post-randomization. Of those randomized, 5.5 percent of respondents attrited from the survey. Robust standard errors are shown. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C10: Patient Change in Beliefs and Trial Representation

| | *Posterior Belief* | | | | *Update Exp. Dir.* | | *Conf. in Beliefs* | |
|---|---|---|---|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Representation | 2.003** | -0.147 | 1.776** | 0.032 | 0.144** | -0.077 | 0.170 | 0.133 |
| | (0.809) | (0.654) | (0.786) | (0.629) | (0.067) | (0.057) | (0.127) | (0.116) |
| Prior on Efficacy | | | 0.105* | 0.109*** | | | | |
| | | | (0.059) | (0.041) | | | | |
| Control Mean | 12.552 | 13.072 | 12.552 | 13.072 | 0.731 | 0.913 | 1.403 | 1.420 |
| Observations | 139 | 136 | 139 | 136 | 139 | 136 | 139 | 136 |

*Notes:* Table reports OLS estimates. In Columns (7)–(8), the *Prior on Efficacy* variable refers to confidence in posteriors on efficacy. *Confidence in Beliefs* is measured on a 1–4 Likert scale, with 4 indicating "High" confidence. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C11: Extrapolation from Clinical Trial Data among Physicians and Patients

| | Not at All (1) | Some (2) | Moderate (3) | High (4) | Perecieved Biol. Factors (5) | Percieved Social & Envir. Factors (6) |
|---|---|---|---|---|---|---|
| **Panel A: Extrapolation from White to Black Patients** | | *Confidence* | | | *Rationale* | |
| Black Patients | 39.6% | 28.1% | 25.2% | 7.2% | 31.0% | 45.7% |
| White Patients | 19.1% | 37.5% | 31.6% | 11.8% | 33.3% | 29.2% |
| PBP | 3.5% | 28.1% | 61.4% | 7.0% | 32.1% | 45.3% |
| PWP | 4.6% | 35.4% | 50.8% | 9.2% | 35.6% | 37.3% |
| **Panel B: Extrapolation from Offshored to U.S. Patients** | | *Confidence* | | | *Rationale* | |
| Black Patients | 33.8% | 34.5% | 22.3% | 9.4% | 21.4% | 54.8% |
| White Patients | 21.3% | 36.8% | 32.4% | 9.6% | 19.5% | 43.9% |
| PBP | 3.5% | 19.3% | 66.7% | 10.5% | 9.8% | 60.8% |
| PWP | 1.5% | 21.5% | 61.5% | 15.4% | 10.9% | 70.9% |

*Notes:* Table reports clinical trial data extrapolation confidence and rationale among patients and physicians. Panel (a) reports confidence in extrapolation across race and Panel (b) reports confidence in extrapolation across geography. Columns (1)–(4) report the percentage of respondents at each confidence level. If a respondent did not select "High" confidence in extrapolation, they were asked to provide a rationale. Column (5) reports the percentage of respondents who cite perceived biol. factors as the rationale for not having "High" confidence in extrapolation. Column (6) reports the percentage of respondents who cite perceived social and envir. factors as the rationale for not having "High" confidence in extrapolation. PBP denotes physicians above the median of reported percent Black population, whereas PWP denotes physicians above the median of reported percent White population. Data are drawn from the Patient Survey Experiment and the Physician Survey Experiment.

Appendix Table C12: Physician Experimental Results — Robustness

| | Relevance Non-Standard | Prescribe Non-Standard | Main Specification | Report Demand Sample | Follow-Up Sample | LASSO Controls |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Representation | 0.109*** | 0.025*** | 0.107*** | 0.121*** | 0.071** | 0.168*** |
| | (0.029) | (0.007) | (0.029) | (0.032) | (0.034) | (0.033) |
| Efficacy | 0.189*** | 1.519*** | 0.281*** | 0.278*** | 0.315*** | 0.224*** |
| | (0.029) | (0.175) | (0.032) | (0.038) | (0.044) | (0.038) |
| Doctor FEs | Yes | Yes | Yes | Yes | Yes | No |
| Profile Order FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Rx Mechanism FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,096 | 1,096 | 1,096 | 784 | 656 | 1,096 |

*Notes:* Table reports OLS estimates. Column (1) reports estimates from Equation 2 with the outcome *Relevance*; in contrast to the primary specification, the outcome is not standardized. Column (2) shows results using *Prescribing Intention*; in contrast to the primary specification, the outcome is not standardized. *Efficacy* is the change in A1C associated with the drug, which has a range of 0.5 to 2.0 percentage points. *Representation* is the share of Black patients in the trial, which ranges from 0 to 35 percent. See main text for a discussion of these ranges. Column (3) reports estimates from Equation 2. Columns (4)–(6) report estimates for the outcome of *Prescribing Intention*, standardized to a mean of 0 and standard deviation of 1. Column (4) shows results restricting the sample to those demanding the NIH/NASEM report. Column (5) displays results restricting the sample to those responding to the follow-up question. Column (6) reports estimates from double-selection LASSO linear regression with potential controls including age, sex, education, and health variables among others. Robust standard errors clustered at the physician level are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C13: Characteristics of Patients Demanding Report

| | All Patients | Demanded Report | Difference Between Groups |
|---|---|---|---|
| | (1) | (2) | (3) |
| Black | 0.505 | 0.548 | -0.042 |
| | (0.501) | (0.500) | (0.056) |
| Male | 0.393 | 0.426 | -0.033 |
| | (0.489) | (0.497) | (0.055) |
| Age Group | 5.876 | 5.870 | 0.007 |
| | (1.117) | (1.166) | (0.126) |
| BA or Higher | 0.287 | 0.287 | 0.000 |
| | (0.453) | (0.454) | (0.050) |
| Insured | 0.931 | 0.948 | -0.017 |
| | (0.254) | (0.223) | (0.027) |
| Takes BP Medication | 0.889 | 0.886 | 0.003 |
| | (0.315) | (0.319) | (0.035) |
| Past Nonadherence | 0.171 | 0.165 | 0.006 |
| | (0.377) | (0.373) | (0.042) |
| General Trust | 0.527 | 0.557 | -0.029 |
| | (0.500) | (0.499) | (0.056) |
| Pharma Trust | 1.636 | 1.730 | -0.094 |
| | (0.801) | (0.798) | (0.089) |
| Doctor Trust | 2.324 | 2.322 | 0.002 |
| | (0.689) | (0.695) | (0.077) |
| Public Health Trust | 1.945 | 2.104 | -0.159* |
| | (0.863) | (0.799) | (0.094) |
| Altruism | 6.793 | 7.357 | -0.564** |
| | (2.188) | (1.812) | (0.231) |
| Risk Preference | 5.422 | 5.861 | -0.439 |
| | (2.516) | (2.509) | (0.279) |
| Time Preference | 6.993 | 7.348 | -0.355 |
| | (1.985) | (2.086) | (0.224) |
| Heard of Tribenzor | 0.047 | 0.043 | 0.004 |
| | (0.213) | (0.205) | (0.023) |
| Prior on Efficacy | 5.782 | 5.696 | 0.086 |
| | (7.131) | (7.514) | (0.805) |
| Observations | 275 | 115 | 390 |

*Notes:* Table compares baseline characteristics of all patient respondents to those demanding the personalized report. Robust standard errors are used when comparing characteristics between the two groups. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C14: Physician Experimental Results — Additional Results

| | Main Specification (1) | Interacted Specification (2) | Continuous Patient Black (3) | Above Median Black (4) | High Black vs. White Strata (5) | High Black vs. Other Two Strata (6) |
|---|---|---|---|---|---|---|
| Representation (Own Group) | 0.107*** | 0.107*** | -0.005 | 0.051 | -0.022 | 0.062 |
| | (0.029) | (0.029) | (0.039) | (0.035) | (0.050) | (0.039) |
| Efficacy | 0.281*** | 0.281*** | 0.285*** | 0.280*** | 0.259*** | 0.302*** |
| | (0.032) | (0.032) | (0.043) | (0.040) | (0.076) | (0.048) |
| Representation × Efficacy | | -0.004 | | | | |
| | | (0.023) | | | | |
| Representation × Patient Percent Black | | | 0.004*** | | | |
| | | | (0.001) | | | |
| Efficacy × Patient Percent Black | | | 0.000 | | | |
| | | | (0.001) | | | |
| Representation × Above Median Black | | | | 0.134** | | |
| | | | | (0.058) | | |
| Efficacy × Above Median Black | | | | 0.006 | | |
| | | | | (0.065) | | |
| Representation × Top Decile Black ZIP | | | | | 0.194*** | 0.107* |
| | | | | | (0.065) | (0.057) |
| Efficacy × Top Decile Black ZIP | | | | | -0.013 | -0.051 |
| | | | | | (0.085) | (0.061) |
| Doctor FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Profile Order FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Rx Mechanism FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,096 | 1,096 | 1,096 | 1,096 | 752 | 1,096 |

*Notes:* Table reports OLS estimates for the outcome of *Prescribing Intentention*, standardized to a mean of 0 and standard deviation of 1. Column (1) reports estimates from Equation 2. Column (2) includes the two-way interaction between *Representation* and *Efficacy* (see Subsection A.2 for further explanation). Column (3) includes interactions between *Representation* and *Patient Percent Black* (reported patient panel percent Black), *Efficacy* and *Patient Percent Black*, and the main effect of patient panel percent Black (the latter is not reported here). Column (4) includes interactions between *Representation* and *Above Median Black* (an indicator for above the median patient panel percent Black), *Efficacy* and this indicator, and the main effect of above the median patient percent Black (the latter is not reported). Column (5) includes interactions between *Representation* and *Top Decile Black ZIP* (an indicator for belonging to a zip code in the top decile of population percent Black), *Efficacy* and this indicator, and the main effect (not reported here) among the sample of physicians either from the top decile or bottom decile of population percent Black. Column (6) reports results from the same specification as Column (5) but estimated among physicians from all three sampling strata (top decile population percent Black, bottom decile population percent Black, and all other deciles). Robust standard errors clustered at the physician level are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C15: Overall Sentiment and Screen Time among Patients

| | *Overall Sentiment* | | *Duration (Min.)* | |
|---|---|---|---|---|
| | Black Patients | White Patients | Black Patients | White Patients |
| | (1) | (2) | (3) | (4) |
| Representative Treatment | 0.039 | 0.180 | -2.672 | 1.750 |
| | (0.114) | (0.112) | (2.758) | (1.783) |
| *p*-value: Black=White | | 0.376 | | 0.176 |
| Constant | 0.045 | 0.014 | 17.595 | 12.522 |
| Observations | 139 | 136 | 139 | 136 |

*Notes:* Columns (1) and (2) report OLS estimates for *Overall Sentiment*, a likert scale from -1 to 1 measuring overall sentiment as implied by open-text responses. Respondents were asked to explain why or why not a given drug was relevant for their own medical care immediately following exposure to the treatment. Continuous sentiment scores (from $-1$ to 1 with 1 being the most positive) are predicted using Valence Aware Dictionary and Sentiment Reasoner (VADER). If the score of an open-text response is greater than or equal to 0.1, a measure of overall sentiment is coded as 1; if the score is less than or equal to $-0.1$, a measure is coded as $-1$; otherwise a measure is coded as 0. Columns (3) and (4) report OLS estimates where the outcome *Duration* is the total time a respondent spent completing the survey in minutes. *Black Patients* denotes the group where the patients self-report their race as Black, and *White Patients* denotes the group where the patients self-report their race as White. *Representatitive Treatment* is an indicator for whether the respondent was assigned to see data from a representative trial (treatment group). To test the null hypotheses that the coefficients for Black and White patients are equal, the *p*-values for both regressions are reported, respectively. Robust standard errors are reported in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C16: Trial Sites at Safety Net Hospitals

| | (1) | (2) | (3) |
|---|---|---|---|
| HIV/AIDS (Cancer Comparison) | 0.117*** | | |
| | (0.008) | | |
| HIV/AIDS (ADRD Comparison) | | 0.150*** | |
| | | (0.012) | |
| Cancer (ADRD Comparison) | | | 0.032*** |
| | | | (0.009) |
| Constant | 0.469*** | 0.436*** | 0.437*** |
| | (0.001) | (0.009) | (0.009) |
| Observations | 196,202 | 6,853 | 195,026 |

*Notes:* Table reports OLS estimates from a regression of an indicator for whether a trial site is at a safety net hospital (SNH). Each observation represents a specific site associated with a unique clinical trial and the data are limited to Cancer, HIV/AIDS, and ADRD trials. Following Popescu et al. (2019), we define an SNH as a hospital in the top quartile within the state it is located in either receiving a disproportionate share of funding from Medicaid or high uncompensated care. In this table, we use the Disproportionate Share Hospital (DSH) Index to define an SNH, drawn from CMS Cost Reports. *HIV/AIDS (Cancer Comparison)* is an indicator variable equal to one if a trial site studies HIV/AIDS and zero if a trial site studies cancer. *HIV/AIDS (ADRD Comparison)* is an indicator variable equal to one if a trial site studies HIV/AIDS and zero if a trial site studies ADRD. *Cancer (ADRD Comparison)* is an indicator variable equal to one if a trial site studies cancer and zero if a trial site studies ADRD. Trial site information is drawn from ClinicalTrials.gov. See Appendix H.1.1 and H.3.8 for details on data. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C17: Neighborhood Demographics of HIV/AIDS, Cancer, and ADRD Study Sites

| | HIV/AIDS | Cancer | ADRD | HIV vs. Cancer | HIV vs. ADRD | Cancer vs. ADRD |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Male | 0.495 | 0.490 | 0.488 | 0.006*** | 0.008*** | 0.002*** |
| | (0.042) | (0.042) | (0.036) | (0.001) | (0.001) | (0.001) |
| Share Under 18 | 0.160 | 0.183 | 0.181 | -0.024*** | -0.021*** | 0.003*** |
| | (0.077) | (0.070) | (0.067) | (0.001) | (0.001) | (0.001) |
| Share 65+ | 0.134 | 0.151 | 0.164 | -0.016*** | -0.030*** | -0.013*** |
| | (0.067) | (0.077) | (0.098) | (0.001) | (0.002) | (0.001) |
| Share Non-Hispanic White | 0.481 | 0.585 | 0.573 | -0.104*** | -0.092*** | 0.013*** |
| | (0.237) | (0.231) | (0.234) | (0.003) | (0.004) | (0.003) |
| Share Non-Hispanic Black | 0.197 | 0.144 | 0.130 | 0.054*** | 0.068*** | 0.014*** |
| | (0.216) | (0.172) | (0.157) | (0.003) | (0.003) | (0.002) |
| Share Hispanic | 0.190 | 0.157 | 0.185 | 0.033*** | 0.005 | -0.028*** |
| | (0.193) | (0.172) | (0.204) | (0.003) | (0.004) | (0.002) |
| Share Non-Hispanic Asian | 0.096 | 0.079 | 0.079 | 0.017*** | 0.017*** | -0.001 |
| | (0.098) | (0.087) | (0.084) | (0.001) | (0.002) | (0.001) |
| Share Non-Hispanic AIAN | 0.003 | 0.004 | 0.003 | -0.001*** | -0.000** | 0.001*** |
| | (0.006) | (0.011) | (0.007) | (0.000) | (0.000) | (0.000) |
| Share Non-Hispanic NHPI | 0.001 | 0.001 | 0.001 | -0.000 | 0.000 | 0.000 |
| | (0.007) | (0.007) | (0.005) | (0.000) | (0.000) | (0.000) |
| Share with Health Insurance | 0.917 | 0.920 | 0.917 | -0.003*** | 0.000 | 0.003*** |
| | (0.060) | (0.058) | (0.059) | (0.001) | (0.001) | (0.001) |
| Share with Private Insurance | 0.680 | 0.710 | 0.708 | -0.030*** | -0.028*** | 0.002 |
| | (0.178) | (0.153) | (0.145) | (0.002) | (0.003) | (0.002) |
| Share with Public Insurance | 0.321 | 0.315 | 0.318 | 0.006*** | 0.003 | -0.003* |
| | (0.145) | (0.125) | (0.127) | (0.002) | (0.003) | (0.002) |
| Share in Poverty | 0.151 | 0.134 | 0.134 | 0.017*** | 0.017*** | 0.000 |
| | (0.049) | (0.046) | (0.041) | (0.001) | (0.001) | (0.001) |
| Share SNAP Recipients | 0.049 | 0.043 | 0.043 | 0.006*** | 0.006*** | -0.001** |
| | (0.023) | (0.021) | (0.020) | (0.000) | (0.000) | (0.000) |
| Observations | 5,606 | 58,963 | 6,948 | | | |
| F-test *p*-value | | | | <0.001*** | <0.001*** | <0.001*** |

*Notes:* Table displays the average demographic characteristics of hospital service areas (HSAs) in which study sites for HIV/AIDS, cancer, and Alzheimer's disease and related dementias (ADRD) are located. HIV/AIDS, Cancer, and ADRD sites represent sites from Clinical Trials.gov. Each site is included once for a given condition. F Test *p*-value is from a regression of the *HIV vs. Cancer*, *HIV vs. ADRD*, and *Cancer vs. ADRD* indicator on all characteristics. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Appendix Table C18: Neighborhood Demographics of HIV/AIDS, Cancer, and ADRD Study Sites with Substantial Federal Investment

| | HIV/AIDS | Cancer | ADRD | HIV vs. Cancer | HIV vs. ADRD | Cancer vs. ADRD |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Male | 0.495 | 0.485 | 0.492 | 0.010* | 0.003 | -0.007 |
| | (0.052) | (0.038) | (0.032) | (0.006) | (0.008) | (0.006) |
| Share Under 18 | 0.148 | 0.130 | 0.159 | 0.018* | -0.011 | -0.029** |
| | (0.083) | (0.076) | (0.076) | (0.010) | (0.013) | (0.013) |
| Share 65+ | 0.122 | 0.112 | 0.148 | 0.011 | -0.026** | -0.037*** |
| | (0.058) | (0.064) | (0.103) | (0.008) | (0.012) | (0.013) |
| Share Non-Hispanic White | 0.458 | 0.513 | 0.481 | -0.054* | -0.023 | 0.031 |
| | (0.234) | (0.209) | (0.238) | (0.029) | (0.039) | (0.037) |
| Share Non-Hispanic Black | 0.219 | 0.170 | 0.184 | 0.049* | 0.036 | -0.014 |
| | (0.227) | (0.195) | (0.185) | (0.027) | (0.036) | (0.033) |
| Share Hispanic | 0.184 | 0.149 | 0.212 | 0.034 | -0.029 | -0.063** |
| | (0.195) | (0.165) | (0.231) | (0.023) | (0.034) | (0.032) |
| Share Non-Hispanic Asian | 0.104 | 0.127 | 0.086 | -0.022* | 0.018 | 0.041*** |
| | (0.092) | (0.091) | (0.075) | (0.012) | (0.015) | (0.015) |
| Share Non-Hispanic AIAN | 0.003 | 0.003 | 0.004 | -0.001 | -0.002* | -0.001 |
| | (0.005) | (0.006) | (0.008) | (0.001) | (0.001) | (0.001) |
| Share Non-Hispanic NHPI | 0.001 | 0.002 | 0.001 | -0.002 | -0.000 | 0.002 |
| | (0.001) | (0.013) | (0.002) | (0.001) | (0.000) | (0.002) |
| Share with Health Insurance | 0.923 | 0.933 | 0.919 | -0.010 | 0.004 | 0.014 |
| | (0.063) | (0.055) | (0.060) | (0.008) | (0.010) | (0.010) |
| Share with Private Insurance | 0.682 | 0.743 | 0.666 | -0.061** | 0.016 | 0.077** |
| | (0.193) | (0.174) | (0.176) | (0.024) | (0.031) | (0.030) |
| Share with Public Insurance | 0.317 | 0.266 | 0.348 | 0.052** | -0.030 | -0.082*** |
| | (0.161) | (0.154) | (0.147) | (0.020) | (0.026) | (0.026) |
| Share in Poverty | 0.162 | 0.146 | 0.147 | 0.017*** | 0.016* | -0.001 |
| | (0.049) | (0.046) | (0.050) | (0.006) | (0.008) | (0.008) |
| Share SNAP Recipients | 0.055 | 0.046 | 0.048 | 0.009*** | 0.007* | -0.002 |
| | (0.024) | (0.024) | (0.025) | (0.003) | (0.004) | (0.004) |
| Observations | 135 | 115 | 54 | | | |
| F-test *p*-value | | | | <0.001*** | 0.168 | <0.001*** |

*Notes:* Table displays the average demographic characteristics of hospital service areas (HSAs) in which study sites for HIV/AIDS, cancer, and Alzheimer's disease and related dementias (ADRD) are located. HIV/AIDS sites represent sites from the HIV Prevention Trials Network (HPTN), HIV Vaccine Trials Network (HVTN), and AIDS Clinical Trials Group (ACTG). Cancer sites represent National Cancer Institute (NCI)-designated cancer centers, National Comprehensive Cancer Network (NCCN) member institutions, and Association of American Cancer Institutes (AACI) cancer centers. ADRD sites represent National Institute on Aging (NIA)-funded Alzheimer's disease research centers. As some sites belong to multiple networks, each site is included once for a given condition. F Test *p*-value is from a regression of the *HIV vs. Cancer*, *HIV vs. ADRD*, and *Cancer vs. ADRD* indicator on all characteristics. See Appendix H.3.3 and H.3.4 for details on data. Robust standard errors are in parentheses. *, **, *** refer to statistical significance at the 10, 5, and 1 percent level, respectively.

# D Selected Quotes

Table D1: Physician Quotes on Extrapolating from the Physician Survey Experiment

| **Extrapolation Across Race** | | |
|---|---|---|
| Subthemes | Selected Quotes | Physician |
| Social and Envir. Factors | "I would feel confident that a White-only sample could demonstrate safety, but I would be less confident about its real-world effectiveness for Black patients, given that Black patients have different lived experiences that result in differential experiences of health care and health outcomes." | PBP |
| | "Race is a social construct, not a biological construct. We risk mistreating Black patients by assuming differences are biological rather than social." | PWP |
| | "Race is a construct based on skin color not necessarily reflective of biology. Differences in efficacy are multifactorial including a variety of patients gives me confidence the drug will help my patients." | PBP |
| Biol. Factors | "Genetic variations exist in certain demographic groups. For instance, a high proportion of those of East Asian descent are fast metabolizers of Plavix." | PBP |
| | "It is unclear if genetic/racial differences would affect the mechanism of action for a particular drug." | PWP |
| | "Other classes of medications have shown variation in effectiveness among different ethnic groups." | PWP |
| | "Drugs may work differently for black and white patients due to genetic differences. The presence or absence of certain genes may affect the efficacy of a drug, and affect the incidence of side effects." | PWP |
| Combination of Factors | "I don't think we understand the connections and interactions between some biological predispositions and some socioeconomic factors, but studies in the past have shown some drugs to work differently in Black and white patients." | PWP |
| | "There may be subtle differences in genetic, biology and social/environmental situations." | PWP |
| Not Confident in Data | "I am more suspicious of findings in research and researchers who exclude certain groups. Including Latino patients." | PBP |
| | "Drugs may work differently in Black and white patients, so it is preferable to have a diverse population in the study. However, it is often necessary to make a decision based on incomplete data." | PWP |
| | "It can not be automatically assumed that drugs will work equally without sufficient data." | PWP |
| | "The external validity of the study is limited if the study demographics don't represent the patient being treated." | PBP |
| Norms/Preferences | "Studies should be done with a broader racial demographic prior to FDA approval." | PBP |
| | "I would prefer if a trial included the population that I treat." | PBP |
| | "Drug trials should account for different races as a confounding variable." | PBP |

| Extrapolation Across Geography | | |
|---|---|---|
| Subthemes | Selected Quotes | Physician |
| Social and Envir. Factors | "Would be confident in safety, but not confident in effectiveness given different lived experiences (social, environmental, cultural) between US and other countries." | PBP |
| | "It depends on what the sampling is again. Different lifestyles, diets, and activity levels can play a part in changing results based on the study demographics. For instance, Rybelsus reported weight loss in a study for DM2 conducted in Japan where the average weight of the patient receiving the therapy is decently less than my average patient with the same condition. They have greater weight loss with my patient population than the predicted amount based on the study." | PWP |
| Biol. Factors | "Depending on racial or ethnic background of the population in the sample not being much in us or outside us." | PBP |
| | "It depends on the race, not the country." | PBP |
| Combination of Factors | "I would need to know racial and demographic/social factors that may influence the drug's effectiveness, etc." | PWP |
| | "It cannot be assumed that drugs will work equally in different populations without appropriate data." | PWP |
| | "As above, I feel there are subtle differences in genetic, biological, as well as social and environmental situations." | PWP |
| Not Confident in Data | "I do not know enough about studies outside the US to be confident in them." | PWP |
| | "I worked on protocols in Thailand and the level of falsifying data by technicians doing the assays was shocking." | PBP |
| | "Other countries may have less stringent measures of safety and efficacy compared to the United States." | PWP |
| | "Limited information is given to know how the study was conducted and on what population." | PBP |
| | "I need to know more about the country and study methods." | PBP |
| | "I would have to evaluate if high standards were upheld." | PBP |
| Dependent on Study Location | "It would depend upon the country. I would expect Western European and Canadian trials to be similar to my particular patient population." | PWP |
| | "It depends where the study was done, and what the population was." | PBP |

*Notes:* Table displays quotes that demonstrate several themes from physician's open-text survey responses in the Physician Survey Experiment. Any physician who did not select that they were "Highly" confident in extrapolating clinical trial data from Black patients to White patients or from patients outside the U.S. to U.S. patients were given a set of potential reasons (see Table V and Appendix Table C11), and then asked to explain why they had chosen a given response. *PBP* (Physicians treating Black Patients) denotes physicians who report above the median percent Black patients in their patient panel. *PWP* (Physicians treating white Patients) is defined similarly with respect to White patients.

Table D2: Inclusive Infrastructure: Intentionality and Community Engagement Quotes from *NASEM (2022)* Report

| Subthemes | Selected Quotes | Role |
|---|---|---|
| Intentional Site Selection | "And so if you want to be inclusive, you need to then think about how many from that population you want to enroll and begin to work towards that goal. That's number one. I think that goes into the framework of intentionality, right? We need to be intentional." | Study Investigator |
| | "You have to want to do it because expediency will kick in that you need to close the study in one year and you want to get those patients enrolled. But I do think if you start to plan from the beginning to have an inclusive group, that's important." | |
| | "I think in some sense the clinics did that for us, like if this is a clinic that largely serves the homeless population downtown and we partner with that clinic, we don't need to do a lot of extra stuff to reach those patients. Making sure those clinics were priorities for us and we did adjust a lot of our approach in working with the clinic." | Study Investigator |
| Physical and Linguistic Access | "They're coming into the clinic like three days a week to get . . . lab samples and that is a lot of driving, that's a lot of time to . . . have to take off work, or have to take away from family. And not all patients are privileged enough to be able to take time off and come to the center every day." | Study Coordinator |
| | "And so travel to centers.. it's a big barrier... so assisting in transportation centers is important if that's required. Remote monitoring is important because I think why bring people back just to check that they're ok when it can be done remotely." | Study Investigator |
| | "There were two language translations that were required in order to do our study... if you don't have those materials prepared and you don't anticipate the need to have those materials a priori, it sort of becomes a self-fulfilling prophecy in that you're not going to accrue well or at all in those populations." | Study Investigator |
| Community Members Inform Protocol | "Get the input of those who are actually working within the communities. . . I think you will come up with a lot of different ways how . . . to diversify their cohort." | Study Coordinator |
| | "We have community advisory boards that are built very early in the process and each site has a different community advisory board because the issues that come up with each geographic location are very different and the communities to serve are very different. . . We try to get a good representation of age and gender and different types of work and the experience in the community." | Study Investigator |
| | "You would go to the community . . . and say 'I have an idea for research. I'd like your opinion on what the community might feel about this. Am I trying to get too many people? What would I need to establish a relationship? How can I help you to help me hire out of the community so that they can have people that are easily accessible to ask questions?" | Study Coordinator |

| Subthemes | Selected Quotes | Role |
|---|---|---|
| Building Trust with Community | "This one community that I'm thinking about has been a little historically suspicious because of bad experiences they've endured of medical research and perhaps academic medical research and so sending out a single notice is not going to be sufficient in order to have meaningful recruitment of these groups. It's really going to start with building relationships of trust and then later availing those groups of opportunities." | Study Investigator |
| | "I would give a talk and try to sit with people. And we had food afterwards usually, so we could all just sit and talk casually. But they're telling me, over and over again, there's just a lot of distrust in the medical community and I get it, I understand why." | Study Investigator |
| Relationships with Community Partners | "And so we do try to give back. We don't just recruit, we always try to give back to the community. I think that's really important if you want to have a relationship with the community, you don't just take. Whatever that community is, we try to teach you, we go to health fairs, we try to give something back." | Study Investigator |
| | "Our academic partners have been working with these community organizations and actually have community health workers who worked with them on other projects. It's easy to take them from one project to another until they have this track record. And it works really nicely for them because they have built in trust already." | Study Investigator |
| | "It's a two-way street. I don't just go to them when I have a study. And I can't expect them to be open and ready to help me with every study and I'm not truly there for them. It's not only me, but it's like having this kind of relationship that is enduring and takes time to build. And it's not a trivial commitment. It's a real long-term commitment. And so we built these relationships with our community partners for now more than a decade and have been and those relationships come with both give and take of information." | Study Investigator |
| Reciprocal Relationships with Participants | "You really can't separate participating in a trial from how a person feels the system treats them. What surprised me is that it cuts across socioeconomic classes. Even my fellow African American physicians express some concerns you would not expect them to express. It's percolating in the back of their minds." | Study Investigator |
| | "Yeah, incentives, we paid them. And then establishing that personal connection with them because they were letting us into their homes with these video recorders and things. I would talk to them on the phone each week. And sometimes these conversations would last 15 minutes, sometimes they would last two hours. Where we would just chat about 'How's it going?' I really tried to get to know them on a personal level." | Study Investigator |

| Subthemes | Selected Quotes | Role |
|---|---|---|
| | "That's why we don't force, everything's voluntary. . . They can withdraw at any time. We make sure that they instill that in anything that we do, no forcing answering questions. Their well-being is first, the study goes second. And then it just always comes first with us because we just put them first. So they put the study first." | Study Investigator |

*Notes:* Table displays quotes from researchers and physicians on inclusive infrastructure policies. Quotes are drawn from NASEM (2022) and STAT News (2020).

# E Survey Instrument

## E.1 Physician Survey Panel

For the Physician Survey Experiment, we recruited 137 physicians from the United States using a physician contact information database from Redi-Data Inc.. We contacted 12,192 primary care physicians to participate in the study, and 1.8 percent of those contacted started the study.[69] Respondents were channeled through demographic questions to ensure that the individual met the pre-specified criteria for our final sample (see Section IV.1.1 for details). The survey was run in March 2022. Respondents were paid only if they fully completed the survey. The compensation for survey completion was $100. The median time for survey completion was 18 minutes.

## E.2 Patient Survey Panel

For the Patient Survey Experiment, we recruited 275 individuals from the United States using a survey company, Lucid. The survey was run in March 2022. Lucid partners with suppliers that provide panels of respondents to which they email survey links.[70] Respondents who clicked on the link were first channeled through demographic questions and questions about High Blood Pressure/Hypertension to ensure that the individual met the pre-specified criteria for our final sample. Respondents were paid only if they fully completed the survey. Respondents were blocked from completing the survey multiple times or reattempting the survey if they were screened out. The pay per survey completed was around $3. The median time for the completion of the survey was 11 minutes.

[69] Due to DUA restrictions and IRB guidance regarding protecting subject privacy, we are not able to release the contact information of the physicians we contacted. See replication materials for de-identified datasets.

[70] See more information on Lucid panels at https://luc.id/quality/

## E.3 Survey Materials

Appendix Exhibit E1: Invitation Email Sent to Physicians

HARVARD Kennedy School
MALCOLM WIENER CENTER
for Social Policy

Dear Dr. PHYSICIAN_NAME,

You have been randomly selected to participate in a study to investigate how physicians use information from clinical trials to treat their patients.

Researchers at Harvard University are conducting this study. The study is funded by an independent research center at Harvard, and is not connected with any pharmaceutical company. This study has been approved by the Institutional Review Board.

Your views are highly valuable and we greatly appreciate your willingness to participate. **As a token of our appreciation, we will give you a $100 honorarium if you pass a few screening questions and complete the survey.**

**Your anonymized views will be used to draft a report to the National Institutes of Health and National Academy of Medicine regarding the types of research that clinicians find most useful for their practice.**

We will also send you a copy of this report, if you would like. Simply click "yes" at the end of the survey to receive it.

This survey includes questions about your background and clinical practice, then asks you to rate 8 hypothetical drugs. All data associated with this survey are located on a secure server at Harvard. The survey takes about 15 minutes to complete.

Please click on the link below to access the survey. The link to the survey will expire in 4 days. Thank you for your help.

**https://harvard.az1.qualtrics.com/jfe/form/SV_898DCxd11ZoL2Rg?Q_DL=HD52bVT9EdDESfV_898DCxd11ZoL2Rg_MLRP_djbzTq2daQrFX9A&Q_CHL=email**

Sincerely,



Marcella Alsan, M.D., M.P.H., Ph.D.
Professor of Public Policy
Harvard University

You may email nikhil_shankar@hks.harvard.edu if you would like more time to complete the survey or if you would like a reminder in 2 days to complete the survey.

Click here to opt out of future emails.

Appendix Exhibit E2: Example Drug Profile




**Drug Name: Afinaglutide**

**Mechanism of Action**: Increases levels of incretin, which enhance glucose-dependent insulin secretion
**Study Type**: Double blind active-comparator control trial
**Drug Efficacy:** Lowers Hemoglobin A1C in patients with poorly controlled diabetes by 1.5%
**Sample Size**: 1500 subjects
**Sample Demographics:** 7% Black, 83% white, 10% other

Link to Survey: https://harvard.az1.qualtrics.com/jfe/form/SV_eEugbM54Kx87Fk2

Appendix Exhibit E3: List of Hypothetical Drugs Shown to Participating Physicians

| **Drug Name** | **Mechanism of Action** |
|---|---|
| Atenaburide | Stimulates insulin secretion from pancreatic beta cells |
| Istapiride | Stimulates insulin secretion from pancreatic beta cells |
| Benzapizide | Stimulates insulin secretion from pancreatic beta cells |
| Islogliptin | Inhibits the enzyme DPP-4 from deactivating incretins that stimulate insulin release |
| Methylgliptin | Inhibits the enzyme DPP-4 from deactivating incretins that stimulate insulin release |
| Dolagliptin | Inhibits the enzyme DPP-4 from deactivating incretins that stimulate insulin release |
| Metaglitazone | Increases insulin sensitivity of fat, muscle, and liver tissue |
| Seraglitazone | Increases insulin sensitivity of fat, muscle, and liver tissue |
| Loraglitazone | Increases insulin sensitivity of fat, muscle, and liver tissue |
| Iscagliflozin | Blocks the protein SGLT2 from absorbing glucose in the kidney, so that it is excreted in urine |
| Paragliflozin | Blocks the protein SGLT2 from absorbing glucose in the kidney, so that it is excreted in urine |
| Sotagliflozin | Blocks the protein SGLT2 from absorbing glucose in the kidney, so that it is excreted in urine |
| Betaglutide | Increases levels of incretin, which enhance glucose-dependent insulin secretion |
| Afinaglutide | Increases levels of incretin, which enhance glucose-dependent insulin secretion |
| Fenaglutide | Increases levels of incretin, which enhance glucose-dependent insulin secretion |

*Notes:* Table shows the names and mechanisms of action of the 15 hypothetical drugs shown in the physician survey. Hypothetical drug names were created by importing the most commonly prescribed diabetes drugs by medication class in the Medical Expenditure Panel Survey (MEPS) to ascertain common drug name suffixes for each class and then replacing the prefixes in these drug names with common generic drug name prefixes as published by the National Library of Medicine. Following this process, a total of 15 hypothetical drug names were used in the survey (3 per medication class). Profiles for all drugs ranged in efficacy from 0.5 percent to 2 percent and in percent Black of trial subjects from 0 percent to 35 percent.

## Appendix Exhibit E4: Physician Survey

Link to Survey: https://harvard.az1.qualtrics.com/jfe/form/SV_eEugbM54Kx87Fk2

- Screeners include:
  1. MD/DO degree
  2. Family practice or internal medicine
  3. <50 percent pediatrics
  4. Currently practice primary care

## Appendix Exhibit E5: Physician Survey Outcomes: Question Wording

| Outcome Name | Question Text | Response Options |
|---|---|---|
| | **Primary Outcomes** | |
| Relevance | • How relevant are the findings from this trial for patients with poorly controlled diabetes in your care? | [On a scale of 0 (Not relevant at all) to 10 (Very relevant)] |
| Prescribe | • Thinking about your patient panel in particular, how likely would you be to prescribe [DRUG NAME] for patients with poorly controlled diabetes in your care? | [On a scale of 0 (Very unlikely to prescribe) to 10 (Very likely to prescribe)] |

# Appendix Exhibit E6: Follow-Up Email Sent to Physicians

HARVARD Kennedy School
MALCOLM WIENER CENTER
for Social Policy

Dear Dr. PHYSICIAN_NAME,

On behalf of our research team, I would like to personally thank you for taking the time to complete our survey on clinical practice.

Based on your responses, I am writing with one follow-up question. **Our research team is planning on donating to a non-profit, the Center for Information and Study on Clinical Research Participation (CISCRP), to support recruitment efforts for clinical trials.** We would like your input on how our donation should be allocated.

CISCRP currently has two initiatives:

- Campaign A, which aims to boost trial participation among the general American public, and
- Campaign B, which focuses on boosting clinical trial participation among Americans from under-represented minority communities.

**For every physician who replies, we will donate $5 to CISCRP.** Of the $5 we donate on your behalf, how much would you like to go to Campaign A and how much would you like to go to Campaign B? Please indicate your choice below.

**I would like the research team's $5 donation to be split in the following manner:**

| | | | | | |
|---|---|---|---|---|---|
| **$0 to Campaign A<br>$5 to Campaign B** | **$1 to Campaign A<br>$4 to Campaign B** | **$2 to Campaign A<br>$3 to Campaign B** | **$3 to Campaign A<br>$2 to Campaign B** | **$4 to Campaign A<br>$1 to Campaign B** | **$5 to Campaign A<br>$0 to Campaign B** |

Thank you so much again for your participation. Please note that responding to the follow-up question is voluntary. If you would like a payment of $5 for your time, please click here. Feel free to contact me at rxmd_study@hks.harvard.edu if you have any questions or feedback on our study.

With warmest regards,



Marcella Alsan, M.D., M.P.H., Ph.D.
Professor of Public Policy
Harvard University

Click here to opt out of future emails.

*Notes:* The order of the campaigns was randomized at the individual level, with the diverse campaign denoted as Campaign A for half the respondents and denoted as Campaign B for the other half of the respondents.

## Appendix Exhibit E7: Patient Survey

Link to Survey: https://harvard.az1.qualtrics.com/jfe/form/SV_eyyp71a6ifHhJf8

- Screeners include:
  1. Race (black/white)
  2. Age (35+)
  3. Hypertension (may have some additional diagnoses but may not select all)
  4. Reasonable value for blood pressure
  5. Never taken survey before
  6. Correctly answer attention question

## Appendix Exhibit E8: Patient Survey Outcomes: Question Wording

| Outcome Name | Question Text | Response Options |
|---|---|---|
| | **Primary Outcome** | |
| Relevance | • How relevant are the findings from this study for treating your high blood pressure? | [On a scale of 0 (Not relevant at all) to 10 (Very relevant)] |
| | **Secondary Outcomes** | |
| Belief Updating | • What millimeters of mercury (mmHg) point reduction in systolic blood pressure would you expect to see if you took the medication? | [On a scale of 0 to 25] |
| Ask Doctor | • Would you be interested in asking your doctor (or other primary care provider) about Tribenzor? | [Yes or No] |

# F Model Appendix

## F.1 Model Details

We assume people update their beliefs from trial data on treatment $T$ according to the following equations, which then enter into $\hat{b}(x_i;h^T)$ analogously to how the prior parameters do in Equation 1:

$$\alpha(x_i;h^T) = \alpha(x_i;h^{T-1}) + \bar{x}_T(x_i) \times k_T \tag{4}$$
$$\beta(x_i;h^T) = \beta(x_i;h^{T-1}) + \bar{x}_T(x_i) \times N_T - \bar{x}_T(x_i) \times k_T \tag{5}$$
$$\alpha(h^T) = \alpha(h^{T-1}) + k_T \tag{6}$$
$$\beta(h^T) = \beta(h^{T-1}) + N_T - k_T. \tag{7}$$

## F.2 Further Results

**Corollary 3.** *Given Proposition 1 and physician-patient dyad's decision making process, the demand for a new medication $d(x_i;h^T)$ is influenced by the reported efficacy and representation of a clinical trial $T$ in the following ways:*

1. $\frac{\partial d(x_i;h^T)}{\partial k_T} > 0$*: demand for a new medication is increasing in the efficacy observed in the clinical trial.*
2. $\frac{\partial d(x_i;h^T)}{\partial \bar{x}_T(x_i)} > 0$*: demand for a new medication is increasing in the representation of patients with similar characteristics.*
3. *Assuming that $F_\varepsilon()$ is convex, $\frac{\partial^2 d(x_i;h^T)}{\partial \bar{x}_T(x_i)^2} < 0$. That is, the degree to which increasing representation in a clinical trial increases demand is decreasing in the existing representation of a patient's group.*

**Corollary 4.** *Let $G_b(h^T) = \hat{b}(x_i = 0;h^T) - \hat{b}(x_i = 1;h^T)$ and $G_d(h^T) = d(x_i = 0;h^T) - d(x_i = 1;h^T)$ be the difference between White and Black patients in perceived benefits and demands for a new medication after observing trial $T$, respectively. Assume that clinical trial $T$ has a higher share of White patients relative to Black patients $\bar{x}(x_i = 0) > \bar{x}(x_i = 1)$ and that the perceived benefits, based solely on the history of trials $h^{T-1}$, for Black patients is less than or equal to that for White patients, then*

1. $G_b(h^T) > 0$ *and* $G_d(h^T) > 0$*: There exists a gap in perceived benefits and demand for novel drugs between White patients and Black patients. White patients have higher perceived benefits and demand relative to Black patients.(Existence of Gaps)*
2. *Increasing Black representation to $\bar{x}'(x_i = 1) > \bar{x}(x_i = 1)$, closes the gap in perceived benefits $G_b(h^T) \geq G_b(h^{T'})$ and demand for novel drugs $G_d(h^T) \geq G_d(h^{T'})$. (Representation Closes Gaps)*

## F.3 Proofs

*Proof of Proposition 1.* Note that

$$\hat{b}(x_i;h^T) = m \times \underbrace{\left(\tilde{b} \times \frac{\alpha(x_i;h^T)}{\alpha(x_i;h^T) + \beta(x_i;h^T)}\right)}_{\text{posterior estimate of } \hat{b} \text{ conditional on } x_i \text{ mattering}} + (1-m) \times \underbrace{\left(\tilde{b} \times \frac{\alpha(h^T)}{\alpha(h^T) + \beta(h^T)}\right)}_{\text{posterior estimate of } \hat{b} \text{ conditional on } x_i \text{ not mattering}},$$

where, by Equations (4)-(7) ,

$$\frac{\alpha(x_i;h^T)}{\alpha(x_i;h^T)+\beta(x_i;h^T)} = \frac{\alpha(x_i;h^{T-1})+\bar{x}_T(x_i)\times k_T}{\alpha(x_i;h^{T-1})+\beta(x_i;h^{T-1})+\bar{x}_T(x_i)\times N_T} \tag{8}$$

$$\frac{\alpha(h^T)}{\alpha(h^T)+\beta(h^T)} = \frac{\alpha(h^{T-1})+k_T}{\alpha(h^{T-1})+\beta(h^{T-1})+N_T}. \tag{9}$$

The three parts of the proposition follow from the fact that the partial derivatives of Equations (8)-(9) with respect to $k_T$ and $\bar{x}_T(x_i)$ are positive, and the second derivatives of these two equations with respect to $\bar{x}_T(x_i)$ are negative. □

*Proof of Corollary 1.* Immediate from Proposition 1 and from demand increasing in perceived benefits.

□

*Proof of Corollary 3.* The demand for a novel drug is $1-F_\varepsilon\left(-\left(\hat{b}(x_i;h^T)-n_T-p_T\right)\right)$. The first two parts of this corollary follow from applying the chain rule and the fact that the demand is increasing in perceived benefits (any c.d.f. is an increasing function) and perceived benefits are increasing in $k_T$ and $\bar{x}_T(x_i)$ (Proposition 1).

To show the third part, note that

$$\frac{\partial^2 d(x_i;h^T)}{\partial \bar{x}_T(x_i)^2} = F_\varepsilon^{'}\left(-\left(\hat{b}(x_i;h^T)-n_T-p_T\right)\right)\frac{\partial^2 \hat{b}(x_i;h^T)}{\partial \bar{x}_T(x_i)^2} - \left(\frac{\partial \hat{b}(x_i;h^T)}{\partial \bar{x}_T(x_i)}\right)^2 F_\varepsilon^{''}\left(-\left(\hat{b}(x_i;h^T)-n_T-p_T\right)\right).$$

By assumption, $F_\varepsilon^{''}(.)\geq 0$; by properties of c.d.fs $F_\varepsilon^{'}(.)>0$; and by proposition 1 $\frac{\partial \hat{b}(x_i;h^T)}{\partial \bar{x}_T(x_i)}>0$ and $\frac{\partial^2 \hat{b}(x_i;h^T)}{\partial \bar{x}_T(x_i)^2}<0$. Plugging these comparative statics into the above equation, it is immediate that $\frac{\partial^2 d(x_i;h^T)}{\partial \bar{x}_T(x_i)^2}<0$. □

*Proof of Corollary 4.* The first part of this corollary follows immediately from Proposition 1 and Corollary 3.

For the second part, note that $\bar{x}(x_i=0)+\bar{x}(x_i=1)=1$. Therefore, an increase in $\bar{x}(x_i=1)$ implies a decrease in $\bar{x}(x_i=0)$. By Proposition 1 and Corollary 3, this implies that demand and perceived benefits increases for Black patients and decreases for White patients when $\bar{x}(x_i=1)$ increases. The gaps $G_b(h^T)$ and $G_d(h^T)$ then decrease by definition. □

*Proof of Proposition 2.* Immediate application of Bayes' rule. □

*Proof of Corollary 2.* Immediate. □

*Proof of Proposition 3.* When the fixed costs $f$ are sufficiently large, then $\bar{x}_T=\bar{x}_T^{sq}$, which is increasing in $\bar{x}_Z$ by Corollary 1 and Proposition 2. □

## F.4 The Value and Cost of Recruiting Strategies to the Firm

For simplicity, suppose each firm only develops a single treatment and there are a continuum with measure one of patient-doctor dyads.[71]

Given recruiting strategy $r$, the value to the firm of the treatment if brought to market is given by

$$v_r = \mathbb{E}\left[\sum_{i=0,1}\sigma(x_i)\times(p-mc)\times d(x_i;h^T)|r\right],$$

where $\sigma(x_i)$ equals the share of people who both have characteristics $x_i$ as well as a condition for which the treatment is indicated, $p$ is the price the firm charges for treatment, $mc$ is the marginal cost of treatment, $d(x_i;h^T)$

[71] This simplifies the analysis by allowing us to abstract from dynamic considerations, where a firm might be concerned that its trial-recruitment decisions today could influence their own recruitment costs in the future. What's important for the analysis is merely that a firm does not internalize *all* the benefits from increasing trial representation — *e.g.*, how this will lower future recruitment costs for *other* firms.

is the demand for treatment at price $p$ among patients with characteristics $x_i$ for whom the treatment is indicated, and $\mathbb{E}[\cdot|r]$ equals the firm's expectation under recruiting strategy $r$.

We assume that firms pay the following cost to increase representation from $\bar{x}_T^{sq}$ to $\bar{x}^r$:

$$\begin{aligned} c_r &= f \times 1(\bar{x}^r \neq \bar{x}_T^{sq}) + h\left(\frac{\bar{x}^r - \bar{x}_T^{sq}}{d(x_i=1;h^T)} \times N - \frac{(1-\bar{x}_T^{sq}) - (1-\bar{x}^r)}{d(x_i=0;h^{T-1})} \times N\right) \\ &= f + h\left((\bar{x}^r - \bar{x}_T^{sq}) \times N \times \frac{d(x_i=0;h^{T-1}) - d(x_i=1;h^{T-1})}{d(x_i=0;h^{T-1}) \times d(x_i=1;h^{T-1})}\right), \end{aligned}$$

where $f \geq 0$ is a fixed cost to deviating from the status-quo recruitment strategy (*e.g.*, due to costs of moving the trial location, setting up a new recruitment infrastructure, etc.) and $h(\cdot)$ is an increasing function that takes as its argument the number of additional patients who need to be targeted to increase Black representation from $\bar{x}_T^{sq}$ to $\bar{x}^r$, holding the overall trial sample fixed at $N$.

Note, there are two channels by which *historical underrepresentation* increases the cost for a firm to enroll a more representative trial participation sample. First, patients with underrepresented characteristics $x_i$ become less likely to participate in the trial when targeted, firms have to reach out to more patients to achieve a given level of representation. Second, the patient pool becomes less representative under the status-quo recruitment strategy over time in the absence of any intervention.

## F.5 Numerical Examples

**Example 1.** For a numerical example, suppose people are certain characteristic $x_i$ matters, so $m = 1$. Suppose further that $\alpha(x_i;h^{T-1}) = \beta(x_i;h^{T-1}) = 100$ for all $x_i$, $\tilde{b} = 100$, $k_T = 750$, and $N_T = 1000$. Then the following table shows the perceived benefit of the treatment for White ($x_i = 0$) and Black ($x_i = 1$) patients as a function of Black trial representation ($\bar{x}$).

| | $\bar{x} = .05$ | $\bar{x} = .1$ | $\bar{x} = .2$ |
|---|---|---|---|
| $\hat{b}(x_i = 0;h^T)$ | 70.65 | 70.45 | 70 |
| $\hat{b}(x_i = 1;h^T)$ | 55 | 58.33 | 62.5 |

There are two noteworthy features of this numerical exercise. First, as seen in the second row of the table, representation significantly matters for the perceived treatment efficacy of groups with low representation. Second, comparing the two rows, trial representation of 80 percent versus 95 percent makes essentially no difference to the perceived benefits of treatment to highly-represented White patients, but trial representation of 20 percent versus 5 percent makes a big difference to the perceived benefits of treatment to poorly-represented Black patients.[72]

Returning to Example 1, modify it slightly to suppose that instead of describing treatment $T$ it describes the most similar past treatment to treatment $T$. Then the table shows how historical representation influences priors for a novel treatment. Indeed, Black-patients' prior expectations for the novel treatment are $\tilde{b}_T \times .625$ if their representation in the previous trial was 20 percent and are only $\tilde{b}_T \times .55$ if their representation in the previous trial was 5 percent. This shows how, even when all groups begin with the same priors for $T - 1 = 0$, beliefs diverge when some groups are systematically more represented in trials than others. This shows how a failure to represent groups in a trial today creates an externality where it's more difficult to represent those groups in a trial tomorrow.

To illustrate, return to the earlier numerical example with $\bar{x}_Z = \bar{x}_Z^a = .1$. Suppose that the non-monetary cost of participating in the trial equals $n = 55$ and that $F_\varepsilon(\cdot)$ is logistic. In this case, if the firm sticks with the status-quo recruitment technology in period $T$ (*e.g.*, because $f$ is sufficiently large), then $\bar{x}_T = \bar{x}_T^{sq} = .06$. This shows that one period of underrepresentation leads Black trial participation to drop by almost half.

[72] Instead of boosting the perceived benefits of treatment to Black patients by doubling representation, fixing the size of the sample, the firm could do so by doubling the size of the sample, fixing Black representation. However, it seems unlikely that the latter would be less expensive to the firm or society, motivating our focus on the former.

# G Institutional Details

## G.1 Policy Efforts to Improve Representation in U.S. Government Sponsored Clinical Trials Research

Federal policies aimed at improving representation in clinical research date back at least five decades, and mostly include regulations targeted at federally-funded researchers. Following the passage of the Civil Rights Act of 1964, the National Institutes of Health (NIH) General Clinical Research Centers added notices to grant applications warning that racial discrimination was illegal. Eventually, all domestic grant applicants to the Department of Health and Human Services (HHS) were required to file an assurance of compliance with Title 6 of the Civil Rights Act of 1964, which prohibits discrimination based on race, color, religion, national origin, or sex in services and establishments that require federal funding.

While Title 6 barred discrimination by sex, federal agencies struggled for decades to balance representation of women in clinical trials with attempts to protect pregnant women and fetuses from potentially harmful exposures to unproven drugs.[73] Healthy males were considered the "norm" in study populations, and many researchers believed that including female participants would confound trial results due to factors such as fluctuations in hormone levels (see Epstein (2007) for an excellent history). In response to concerns about these potential harms, the FDA distributed new guidelines in 1977, "General considerations for the clinical evaluations of drugs," that banned women of childbearing potential from Phase I and early Phase II trials with limited exceptions. As documented in Michelman and Msall (2021), this policy change reduced investment into drugs aimed toward female patients. This policy was rescinded in 1993.

In 1985, the U.S. Public Health Service Task Force on Women's Health Issues released a report indicating that low representation of women in clinical trials had resulted in suboptimal health care for women. The task force recommended increasing participation of women in clinical trials, including women of child-bearing potential, and also advised that more research be supported on diseases with a high prevalence among women.

The NIH responded to the taskforce's report by adopting the Inclusion of Women and Minorities in Clinical Research policy in 1986. Although broadly intended to provide information on differences in drug safety and efficacy by sex and race/ethnicity, its implementation was slow and incomplete. Guidelines were developed in 1989, but the resulting requirements for researchers were inconsistently applied. In 1990, the NIH founded the Office for Research on Women's Health (ORWH) and Office of Minority Programs (OMP) in 1990.

The FDA responded to the U.S. Public Health Service's report in 1987 with new guidelines encouraging new drug sponsors to use animals of both sexes in pre-clinical drug safety studies. The following year, the FDA released guidance in which it recommended analyzing data from clinical pharmacology studies by sex, race, and age. These recommendations, however, were not binding.

The 1993 NIH Revitalization Act included additional reforms. Designed to ensure that clinical research could determine differential effects of interventions by sex and race/ethnicity, the Revitalization Act included the following stipulations: Phase III clinical trials should have sufficient numbers of participants to allow for subgroup analyses, populations should not to be excluded from trials due to cost, and the NIH must maintain outreach efforts to include women and minority populations in trials. In 1994, the FDA Office of Women's Health was established to coordinate policies regarding the inclusion of women in clinical trials.

In 1997, Congress enacted regulation requiring the FDA and NIH to review and develop guidance on the inclusion of women and minorities in clinical trials. To comply, the FDA issued the demographic rule in 1998, which revised New Drug Applications (NDAs) to require safety and efficacy data presented by gender, age, and racial subgroups and dosage modifications identified for specific subgroups. In contrast to the non-binding 1980s guidance, the rule gave the FDA the authority to refuse any NDA that did not appropriately analyze safety and efficacy data and applied to all sponsors, regardless of federal funding. Additionally, the demographic rule required sponsored to present data on participation in Investigational New Drug (IND) applications by sex, age, and race. Congress passed a law in 2000 that permitted the FDA to place clinical holds on IND studies if men or women were

[73] In the early 1960s, maternal exposure to the sedative thalidomide during pregnancy led to widespread fetal death and birth defects across Europe and Canada. Congress responded by passing the Kefauver- Harris Amendments in 1962, which strengthened the FDA's authority to oversee drug development and testing.

excluded from clinical trials studying a serious or life-threatening illness on the grounds of threats to reproductive potential.

Under Section 907 of the FDA Safety and Innovation Act of 2012, the FDA issued guidance in 2014 outlining rules regarding sex-specific patient enrollment, data analysis, and reporting of study information. In 2015 the FDA launched the FDA Drug Trials Snapshots database, which provides information about the populations participating in clinical trials associated with new drug applications. FDA Snapshots data must report whether differences in benefits or side effects were detected by sex, race, ethnicity, or age.

### G.1.1 Efforts to Expand Reach of Clinical Trial Recruitment

Cancer research centers have recently implemented a variety of initiatives aimed at boosting recruitment from underrepresented communities; see, for example, https://ncorp.cancer.gov/about/ and https://www.cancer.gov/about-nci/organization/crchd for details. The NIH has recently expanded research networks for ADRD to address low represesentation of Black and Hispanic populations in trials; see https://www.nia.nih.gov/news/nih-expands-alzheimers-and-related-dementias-centers-research-network for details on trial sites in North Carolina and Texas. The CDC's National Healthy Brain Initiative Road Map Series has, in parallel, focused on building partnerships with state and local public health agencies to support ADRD efforts in communities https://www.cdc.gov/aging/healthybrain/roadmap.htm.

### G.1.2 Participant Compensation

In light of the facts documented in Table VII and Appendix Figure B17, extending participation incentives to offset the personal cost of participating in trials is especially important. Transportation subsidies, stipends to offset lost wages and childcare costs, and guaranteed coverage of any supplemental medical care not included in the trial can ensure that geographically distant trials remain accessible to patients (NASEM 2022).[74]

## G.2 Background on ClinicalTrials.gov

ClinicalTrials.gov, a registry of clinical trials maintained by the U.S. National Library of Medicine, was established in 2000. Under the Food and Drug Administration Modernization Act of 1997 (FDAMA), the FDA and NIH were directed to develop an online registry of clinical trials that contained details on all drug efficacy studies associated with approved Investigational New Drug (IND) applications (Food and Drug Administration 2022 Sections 312 & 812). A precursor to ClinicalTrials.gov, the AIDS Clinical Trials Information Services (ACTIS), provided a template.[75]

At its inception, the registry included only a small minority of clinical trials performed domestically and worldwide, with the database largely comprised of NIH-sponsored trials. The completeness of the registry, however, grew over time with the introduction of additional federal and non-governmental reforms. In 2005, the International Committee of Medical Journal Editors (ICMJE) began to require trial registration as a condition of publication, motivating many academics to become careful about registering their trials (ICMJE 2022). Two years later, the FDA Amendments Act (FDAAA) mandated registration and results reporting on ClinicalTrials.gov for all trials studying FDA-regulated drugs. The law authorizes penalties of up to $10,000 per day and the current or future withholding of federal grant funds for violating these provisions (Food and Drug Administration 2020). Following the FDAAA, registration of trials substantially increased.

[74]Most of the literature on improving representation and incentives revolves around ensuring that time and monetary costs are not an undue burden that disproportionately deters certain low socioeconomic groups from participating. Although there is some concern that incentives may still carry the possibility of undue influence, a recent NASEM report clarifies that undue influence occurs when individuals take actions that are not reasonable (NASEM 2022, p.∼100-101). Participation in a trial should always be reasonable for the targeted individual, or else the trial should not pass ethical review. Additionally, incentives should not be used with the goal of "changing the minds" of otherwise hesitant participants.

[75]Consistent with the discussion in our final section, ACTIS was founded under pressure from HIV/AIDS activists, who demanded greater transparency in research.

The FDA's efforts to increase clinical trial diversity and trial transparency grew interrelated over the most recent decade. In 2010, the Affordable Care Act (ACA) required the collection and reporting of demographic data of clinical trial participants; however, a subsequent review by the FDA demonstrated many inconsistencies in industry's adherence to these requirements, particularly with respect to race reporting. In response to these deficiencies, a new FDA rule came into effect in 2017 requiring the submission of "baseline or demographic characteristics measured in the clinical trial, including age, sex/gender, race, [and] ethnicity" for FDA-regulated drugs, if such demographics were collected under the protocol. The enactment of this rule resulted in a large rise in the share of trials reporting race/ethnicity data.

Appendix Exhibit G1: Timeline of Federal Initiatives to Increase Representativeness of Clinical Trials

| Year | Action |
|---|---|
| 1965 | NIH General Clinical Research Centers add notices to grant applications warning that racial discrimination is illegal. |
| 1986 | NIH adopts the Inclusion of Women and Minorities in Clinical Research policy, which is aimed at ensuring that clinical trials are designed to provide information about sex and race/ethnicity differences, but the policy is slow to be implemented and inconsistently applied. |
| 1988 | New FDA guideline recommends analyzing data from clinical pharmacology studies for safety and efficacy by sex, race, and age. |
| 1993 | NIH Revitalization Act directs the NIH to establish guidelines for the inclusion of women and minorities in clinical research. |
| | FDA withdraws policy banning women of childbearing potential from participating Phase I and early Phase II trials, except in the case of life-threatening conditions and with certain other exceptions; the policy had been in place since 1977. |
| 1993 | FDA Office of Women's Health is established to guide the agency around policies on the inclusion of women in clinical trials. |
| 1998 | FDA Demographic rule revises new drug application (NDA) content to require safety and efficacy data by gender, age, and racial subgroups. |
| 2000 | ClinicalTrials.gov, an online registry of clinical trials co-developed by FDA and NIH, is established. The registry initially includes information largely only on NIH-sponsored trials. |
| 2005 | International Committee of Medical Journal Editors (ICMJE) requires trial registration as a condition of publication, raising the number of academics registering on ClinicalTrials.gov. |
| 2007 | FDA Amendments Act introduces reporting requirements for FDA-approved products and penalties for failure to comply (including withholding of federal grant funding and fines of up to $10,000 per day). |
| 2010 | Affordable Care Act (ACA) requires the collection and reporting of demographic data for clinical trial participants. |
| 2015 | FDA's Five Year Plan aims to make pivotal trials (efficacy studies used as basis for new drug approval) more representative of U.S. population and starts publishing of data on composition (the FDA Drug Snapshots). |
| 2017 | FDA Reauthorization Act (FDARA) made a reference encouraging the "enrollment of more diverse patient populations." |
| | FDA rule requiring the submission of age, sex/race, and ethnicity data for trial subjects (if collected) on ClinicalTrials.gov goes into effect. |
| | Requirement for Drug Snapshots to provide information about the trial populations and differences in efficacy/side effects by sex, race, ethnicity, and age. |
| 2020 | FDA releases non-binding, industry-focused guidance on enhancing the diversity of clinical trial populations and that calls for broadening eligibility criteria and improving trial recruitment. |

*Notes:* Table lists major policy initiatives by the FDA, NIH, and HHS to increase diversity and improve representativeness of clinical trials.

# H Data

## H.1 Clinical Trials

We draw on two publicly available sources of clinical trials data: ClinicalTrials.gov and the FDA's Drug Trials Snapshots database.

### H.1.1 ClinicalTrials.gov

We collected clinical trials records from the Aggregate Analysis of ClinicalTrials.gov (AACT) data, a publicly available relational database that contains information on all studies registered on ClinicalTrials.gov (Tasneem et al. 2012).[76] In our analyses, we used the `baseline-counts`, `baseline-measurements`, `countries`, `facilities`, `funding`, and `studies` files. It is useful to note that — to the best of our knowledge — many data files in ClinicalTrials.gov have not been extensively used by researchers and that the nature of the dataset means that our sample is prone to some error.

Many trials in ClinicalTrials.gov are missing data on demographic composition and trial outcomes. In some cases, this is due to changes in reporting requirements over time: between 2008 and 2017, only trials that were conducted under an ever-approved IND — that is, trials that supported either a successful U.S. drug application or U.S.-based marketing for an approved product — were required to report results to ClinicalTrials.gov. After 2017, all trials conducted under an IND registered with the FDA were subject to reporting requirements (Food and Drug Administration 2022). In other cases, missing data reflects widespread noncompliance with reporting requirements.[77]

First, we narrowed our sample to trials that were reported as "completed" to ensure that demographic composition and enrollment statistics were finalized. Of the 393,263 trials in the `studies` file, 211,674 trials were completed. Then, we selected trials that reported sex and race enrollment information, generating 42,005 and 17,612 trials, respectively.

We then classified the clinical trials by disease category using Medical Subject Headings (MeSH) associated with each trial. We grouped MeSH tree numbers and MeSH unique IDs to create disease categories by searching the National Library of Medicine's MeSH search tool, and merging this information with the `browse-conditions` file.[78] We selected disease categories of interest based on the top ten causes of death in the U.S. in 2019 -– diseases of the heart, cancer, chronic lower respiratory diseases, cerebrovascular diseases, Alzheimer's disease, diabetes mellitus, kidney diseases, and influenza and pneumonia (Heron 2021). We also chose to include HIV/AIDS, as it presents an example of well-represented clinical trials (see Section VI.1 for details).[79] We use these data to create Figure VI, and Appendix Figures B1, B2, B3, B13, and B18.

Additionally, we use data from ClinicalTrials.gov for our analysis of trial sites (in, for example, Table VII). We used the `facilities` file and restricted the sample to sites located in the United States. We then restricted our sample to sites for trials studying HIV/AIDS, ADRD, or cancer. We used these data to produce Table VII; Appendix Figure B17; and Appendix Tables C16, C17, and C18. To compute the share of trials with primary sponsors from government and industry, as reported in Section II, we used the `funding` file. We then restricted our sample to U.S.-based trials using the `countries` file with a single leading sponsor.

### H.1.2 FDA Drug Trials Snapshots

In contrast to ClinicalTrials.gov, the FDA Drug Trials Snapshots database reports complete information on the demographic composition of pivotal trials associated with new drug applications approved by the FDA. Per the

[76] We downloaded a version of this database on 15 October, 2021. Dataset versions are released frequently and changes in compliance rules for ClinicalTrials.gov mean that historical records *may* change between versions.

[77] See Anderson et al. (2015) and Zarin, Goodman and Kimmelman (2019) for two discussions of missing data in ClinicalTrials.gov.

[78] MeSH Search Tool available at: https://meshb.nlm.nih.gov/. The 2021 MeSH tree number data can be downloaded from: https://nlmpubs.nlm.nih.gov/projects/mesh/2021/meshtrees/.

[79] We excluded accidents and intentional self-harm from the sample because there are no prescription medicines indicated for these causes of death. We selected 2019 as the reference year in order to exclude Covid-19 deaths.

FDA, the goal of this database is to provide data on clinical trial evidence to patients and is "part of an overall FDA effort to make demographic data [associated with trials] more available and more transparent."[80] A standard entry corresponds to a drug and includes the following pieces of information:

- Drug name and sponsor
- Drug approval date
- Approved indications and method of use
- Any differences in clinical trial evidence by sex, race, or age
- Side effects
- Demographic composition of trials by age, sex, race, ethnicity, location
- Trial design

Beginning in 2015, the FDA's Center for Drug Evaluation and Research (CDER) has published a document containing these data for all new drug approvals. We collected PDFs of Drug Trials Snapshots summary reports for each available year (2015–2021) and digitized tables containing demographic data on trials.[81] For each record, we collected notes about approved indications at the time of marketing approval.

Although FDA Snapshots data include information on the number of trials associated with each drug approval for additional details., it aggregates data to the drug-level rather than the trial-level.[82] That is, FDA Snapshots data indicate the *total* share of Black and White patients enrolled across all trials, but not trial-specific enrollment figures.

A key advantage of using the FDA Drug Trials Snapshots database is that information on the demographic composition of trials is nearly complete. The share of Black patients is missing for 20 of 383 observations. Of these trials missing Black patient share, the enrollment shares of other races sum to 100 percent for six trials, implying that there are no Black participants in these trials. Of the remaining 14 trials, 11 report no race information and three have incomplete race information. We use these data to produce Figure VI and Appendix Figures B1 and B2.

## H.2 Prescribing Rates

We constructed prescribing rates for new drugs (in, for example, Figure I) approved by the FDA by drawing on administrative FDA records and the Medical Expenditure Panel Survey (MEPS). From MEPS, we collected two sets of data for the years 1996 - 2015: information on each survey respondent, including medical histories and demographic information, and information on all prescriptions received by MEPS respondents in the preceding year.[83] The full survey file is nationally-representative with person-weights, which we include in all analyses. MEPS Prescribed Medicines files include the following information: unique MEPS respondent identifier, medication name (imputed), National Drug Code (imputed), a unique identifier for each unique prescription dispensed, Clinical Classification Software (CCS) code (imputed), and prescription year. From the FDA, we used a dataset of new drug applications, restricting to New Drug Applications (NDAs) and Abbreviated New Drug Applications (ANDAs).[84] We will call this "FDA product data" for consistency. We collected National Drug Code, marketing category, and marketing start date from this data.

[80] See https://www.fda.gov/media/97210/download for additional details.

[81] Reports were downloaded from the FDA on 11 Jan 2022: https://www.fda.gov/drugs/drug-approvals-and-databases/drug-trials-snapshots, and data were extracted using OCR tools.

[82] For example, "*The FDA approved ADLYXIN primarily based on evidence from nine clinical trials of 4,508 patients with type 2 DM. The trials were conducted in the United States, Canada, Europe, Australia, South America, Africa, and Asia. The FDA also considered data from one separate trial of 6,068 patients with type 2 DM who recently suffered heart attack. The trials were conducted in the United States, Canada, Europe, Africa, and Asia.*" See https://www.fda.gov/drugs/drug-approvals-and-databases/drug-trials-snapshots-adlyxin

[83] MEPS demographic data were downloaded from IPUMS. Demographic data for 2019 is used to produce Appendix Tables C5 and C6. Data can be downloaded from: https://meps.ipums.org/meps/. MEPS prescription files were downloaded from the Agency for Health Care Policy and Research. Data can be downloaded from: https://www.ahrq.gov/data/meps.html.

[84] These data are current as of September 10, 2022. Data can be downloaded from: https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory.

The goal of our data construction was to link each MEPS prescription record to the marketing start date of the prescribed product. This linkage allowed us to measure the relative age of a prescribed drug (*i.e.*, the number of years between when the manufacturer first marketed the drug in the U.S. and when the patient started taking the drug). Of the 114,624 applications in the FDA product data, we narrowed our sample to 52,562 NDA and ANDA product applications with unique NDCs. Of these 52,562 new drugs, we successfully link 9,895 ($\sim$19%) to at least one prescription record in MEPS. For several reasons, we expect such a match to be imperfect. First, in our read, many drugs in the FDA product data are almost never prescribed to U.S. patients, as they are used to treat conditions that are rarely, if ever, treated in the U.S. (*e.g.*, Ebola and malaria). Additionally, many drugs that aid in clinical imaging or lab analysis are approved as NDAs and ANDAs . Finally, many drugs are dispensed in hospital settings (*e.g.*, many chemotherapy drugs) and do not appear in the MEPS set. Thus, we expect that not all products will be associated with MEPS prescribing records.

We linked the MEPS data to FDA product data using National Drug Codes (NDC) – unique 11-digit identifiers for drugs assigned by the FDA.[85] In both the MEPS and FDA product data, NDCs are reported at the manufacturer-product-package level. Consequently, there are multiple NDCs listed for each product based on the manufacturer, dosage, formula, package size and package type. For example, in the FDA product data, there are 13 unique NDCs for the heart failure/hypertension drug "rivaroxaban." On the other hand, there are only two unique NDCs for "rivaroxaban" in the MEPS data. When we merge the two data sets, only one code matches. Clearly, we are losing some observations due to the discontinuity of NDCs when we perform this merge. Additionally, NDCs are reported only for a subset of observations in the MEPS data, also limiting our sample (363,947 of 5,835,894 observations report NDC code). Ultimately, of the 5,835,894 observations in the MEPS sample, 1,057,031 observations successfully matched to product information in the FDA product data.

After the merge, we dropped observations for which the relative year was negative (*i.e.*, the patient reported that they started taking the drug before the marketing start year). We hypothesize that these inconsistencies may arise if a patient was prescribed an earlier-approved version of the drug than reported in MEPS. However, we dropped these observations because we could not decisively identify the cause. Finally, we dropped observations to drugs with a relative year greater than five, restricting to "new drugs" and yielding 357,312 observations.

Next, we used CCS codes to group the MEPS observations by diagnosis, allowing us to calculate the prescription rates across disease categories (in, for example, Figure VI). CCS codes collapse diagnosis and procedure codes from the International Classification of Diseases, 9th Revision, Clinical Modification (IDC-9-CM), which contains over 14,000 diagnosis codes and 3,500 procedure codes, into 260 disease categories.[86] We restricted our MEPS sample to 1996–2015, because years after 2015 do not report ICD-9 or CCS code. We selected disease categories of interest based on the top ten causes of death in the U.S. in 2019– diseases of the heart, cancer, chronic lower respiratory diseases, cerebrovascular diseases, Alzheimer's disease, diabetes mellitus, kidney diseases, and influenza and pneumonia (Heron 2021). HIV/AIDS is also included as it presents an example of well-represented clinical trials (see Section VI.1).

By nature of the MEPS prescribing data, we recognize that the CCS code variable does not directly map to the indicated use of the prescription drug. For example, a patient may be diagnosed with diabetes mellitus, but be prescribed an antidepressant. In the MEPS data, the observation will report the NDC for the antidepressant, but the CCS code for diabetes mellitus. Hence, the antidepressant will appear in the diabetes mellitus disease category. We recognize this limitation of our classification system, but find that the drugs that appear with the highest frequency appropriately map to the disease category. Furthermore, drugs not associated with the disease or its comorbidities generally do not significantly vary across disease categories, and drugs associated with the disease or its comorbidities are those we would be interested in including. In fact, we specified that the cancer category is cancer supportive therapies, because the MEPS prescriptions mapped to cancer largely represent drugs that treat symptoms and comorbidities associated with cancer. A large majority of cancer therapies are prescribed and administered in the inpatient and office setting and are not captured in the MEPS data.

Though we do not consider our matching process perfect, we are unable to identify a more comprehensive approach in the literature. Using NDCs as a crosswalk provides a streamlined and reliable link between MEPS and product information. An alternative approach would be to link MEPS observations to FDA product data by active

[85]The NDCs are in different formats in MEPS and FDA product data sets, and thus required cleaning. We followed an approach similar to Roth 2018 to create a crosswalk.

[86]See more details on CCS codes at https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp.

ingredient and/or product name, but this method is subject to error as well -– the complexity and nonconformity of active ingredients and product naming conventions across data sets would lead to false matches.

This process yielded our primary data set at the patient-drug-prescription year level, which contains 357,312 observations corresponding to 3,509 products, 44 prescription years, and 64,015 distinct MEPS respondents. We use these data to produce Figures I and VI; and Appendix Figures B3, B4, and B18.

## H.3 Other Data Sources

### H.3.1 Physician Information

Data published by the American Medical Association (AMA) and distributed in our case by Medical Marketing Service Inc., provide a host of information at the physician-level such as location of birth, medical school code and year of graduation, location of practice, and specialty. We calculate average physician characteristics at the zip code-level.[87]

Next, we used National Plan and Provider Enumeration System (NPPES) National Provider Identifier (NPI) Registry data, published by the Centers for Medicare & Medicaid Services (CMS), to link health care providers with their NPIs. This crosswalk linked to our survey based on physician name and to the AMA Masterfile based on NPI.[88] For physicians in our survey data not matched to an NPI using the NPPES dataset, we attempted to manually match them based on name and practice location using the NPPES's online tool.[89]

We accessed zip code-level counts of physicians from Redi-Data Inc. We requested counts of physicians who are 1) aged 30-70; 2) hold a DO or MD degree; and 3) actively practice primary care in an office setting.[90] We use data in this section to produce Appendix Table C3.

### H.3.2 U.S. Medical Schools

We obtained data on U.S. medical school codes from the Texas Medical Board website and created a data set using Adobe Acrobat converter.[91] Additionally, we used rankings from U.S. News and World Report 2022 medical school research rankings, and manually created a data set.[92] We use these data to produce Table C3.

### H.3.3 U.S. Census

The 2019 American Community Survey (ACS) provides zip code-level population and socioeconomic demographic 5-year estimates.[93] [94] We use these data to produce Appendix Figure B17, Appendix Tables C3, C17, and C18. We obtained estimates of the U.S. Black population share and White population share from the U.S. Census website (U.S. Census Bureau 2021). We use these data to produce Figure I and Appendix Figure B1.

### H.3.4 Clinical Trials Network Locations

We built a data set of clinical trial network site locations using information gathered from multiple network websites for cancer, HIV/AIDS, and ADRD. We collected data from the National Cancer Institutes (NCI), National

[87]These data are DUA-restricted and we are thus unable to release them to the public. The authors will assist with any reasonable replication attempts for two years following publication.

[88]Data available at https://download.cms.gov/nppes/NPI_Files.html.

[89]Available at https://npiregistry.cms.hhs.gov/. We are not able to release the dataset of manual matches because NPI numbers are personally identifiable and would enable individuals to identify the identity of our subjects.

[90]See replication file for more detail on Redi-Data access.

[91]Data available at https://www.tmb.state.tx.us/idl/1A8F42C3-00C3-5F24-D581-61051CA584BC.

[92]Medical schools given a ranking ranging between two numbers in U.S. News were assigned the midpoint of those two numbers in the data set, and those that were unranked were assigned a ranking of 124 as the U.S. News rankings stop at 124, following the methodology in Alsan et al. (2019). Data available at https://www.usnews.com/best-graduate-schools/top-medical-schools.

[93]Data available at https://data.census.gov/cedsci/table?q=DP03&g=0100000US%248600000&tid=ACSDP5Y2020.DP03.

[94]Data available at https://data.census.gov/cedsci/table?q=DP05&g=0100000US%248600000&tid=ACSDP5Y2020.DP05.

Comprehensive Cancer Network (NCCN), and Association of American Cancer Institutes (AACI) websites to build a database of cancer network sites.[95] We collected data from the HIV Prevention Trial Network (HPTN), AIDS Clinical Trial Group (ACTG), and HIV Vaccine Trials Network (HVTN) websites to build a database of HIV/AIDS network sites.[96] We collected data from the NIA Alzheimer's Disease Research Center (ADRC) to build a database of ADRD network sites.[97] We use these data to produce Appendix Table C16 and Appendix Table C18.

### H.3.5 Research!America

The nonprofit Research!America fields frequent surveys to gauge public opinion on attitudes toward medical, health, and scientific research. We use data from three survey waves: 2013, 2017, and 2021. We constructed and recoded six variables using the Research!America data (See Appendix Table H2). We simply recoded the responses to create a uniform scale across Variables (1)–(5). A value of one corresponds with the most negative response while a value of four corresponds with the most positive response. As a result, we were able to construct the Kling-Liebman-Katz Index (Kling, Liebman and Katz 2007) in Table I.

We constructed Variable (6) by combining two questions from the Research!America survey — "In general, do you think the work that scientists do benefits all, most, or very few people in the country?" and "In general, to what extent do you think the work that scientists do benefits you?" If a respondent answered "All" or "Most" to the former question, or "To a great extent" or "Somewhat" to the latter question, we assigned "Yes" to the Variable (6). We use these data to produce Table I and Appendix Table C1.

[95] NCI data available at https://www.cancer.gov/research/infrastructure/cancer-centers/find, NCCN data available at https://www.nccn.org/home/member-institutions/, AACI data available at https://www.aaci-cancer.org/members.

[96] HPTN data available at https://www.hptn.org/research/sites, ACTG data available at https://actgnetwork.org/sites-list/, and HVTN data available at https://www.hvtn.org/participate/study-clinics.html#north-america.

[97] ADRC data available at https://www.nia.nih.gov/health/alzheimers-disease-research-centers.

Appendix Exhibit H2: Research!America Codebook

| | Variable | Question | Raw Responses | Recoded Responses | Years | Black | White |
|---|---|---|---|---|---|---|---|
| (1) | Confidence in Research Institutions | "How confident are you in research institutions?" | 1: A great deal<br>2: Some<br>3: Not much<br>4: None at all<br>5: Not sure | 1: None at all<br>2: Not much<br>3: Some<br>4: A great deal<br>Missing: Not sure | 2021 | 215 | 815 |
| (2) | Heard of Clinical Trial | "Have you ever heard of a clinical trial?" | 1: Yes<br>2: No<br>3: Not sure | 0: No<br>1: Yes<br>Missing: Not sure | 2013<br>2017<br>2021 | 659 | 2184 |
| (3) | Would Enroll if Doctor Recommends | "If your doctor found a clinical trial for you and recommended you join, how likely would you be to participate in a clinical trial?" | 1: Very likely<br>2: Somewhat likely<br>3: Not likely<br>4: Would not participate<br>5: Not sure | 1: Would not participate<br>2: Not likely<br>3: Somewhat likely<br>4: Very likely<br>Missing: Not sure | 2013<br>2017<br>2021 | 637 | 2021 |
| (4) | Trust Not Reason for Lack of Enrollment | "Which of the following do you think is a reason that individuals don't participate in clinical trials?" | 0: Not select lack of trust<br>1: Selected lack of trust | 0: Selected lack of trust<br>1: Not select lack of trust | 2013<br>2017 | 581 | 1450 |
| (5) | Would Get FDA-Approved Vaccine | "What is the likelihood of getting COVID vaccine if it was FDA-approved?" | 1: Very likely<br>2: Somewhat likely<br>3: Not very likely<br>5: Not sure | 4: Very likely<br>3: Somewhat likely<br>2: Not likely<br>Missing: Not sure | 2021 | 121 | 801 |
| (6) | Science is Beneficial | "In general, do you think the work that scientists do benefits all, most, or very few people in this country?"<br>—<br>"In general, to what extent do you think the work that scientists do benefits you?" | 1: All<br>2: Most<br>3: Some<br>4: Very few<br>5: Not sure<br>—<br>1: To a great extent<br>2: Somewhat<br>3: A little<br>5: Not sure | 0: No<br>1: Yes | 2021 | 134 | 837 |

*Notes:* Table lists variables constructed using Research!America data, the corresponding question, the response options, the years the question was asked, and the number of observations for Black and White respondents.

### H.3.6 Moderna Enrollment

We construct data on enrollment in trials run by Moderna Therapeutics Inc. associated with the novel coronavirus vaccine. To do so, we collected records of enrollment numbers from public presentations and announcements by firm executives. Data on Moderna's Phase III trial enrollment target (30,000 participants) were drawn from National Institutes of Health (2020) and were publicly announced by Chief Executive Officer Stéphane Bancel. Data on the weekly racial composition of new enrollees were sourced from a presentation by Chief Development Officer Melanie Ivarrson to the National Academies of Sciences Engineering and Medicine on March 29, 2021 (Ivarsson 2021). Data on the weekly accrual rate are drawn from a figure published on the Moderna Therapeutics Inc. website. The figure has been removed from the website, so we used the "Wayback Machine" to view the figure (Moderna Therapeutics 2020). We extracted data from the figure to estimate daily participant recruitment. Note that the accrual rates in Figure B5 Panel (c) are close estimates, as we were not able to source the exact number of participants recruited each day. Data in this section are used to produce Figure B5.

### H.3.7 Stock Prices

Data on the Moderna Therapeutics Inc. (MRNA) stock price are sourced from the Yahoo!Finance database (Yahoo!Finance Historical Data 2022).[98] We obtained the closing price and volume for the dates July 27, 2020–October 19, 2020. These data are used to produce Figure B5.

### H.3.8 Safety Net Hospitals

Definitions of safety net hospital are from Popescu et al. (2019). The first definition is based on the cost of uncompensated care. Uncompensated care is defined by the cost of charity care plus non-Medicare and non-reimbursable Medicare bad debt. We computed uncompensated care using Centers for Medicare & Medicaid Services (2019), American Hospital Association's 2020 annual survey to calculate the cost of uncompensated care as a percentage of total operating expenses.[99] The second definition of a safety net hospital relies on Medicaid Disproportionate Share (DSH) Index. The DSH index is defined by the CMS Cost Reports as the number of Medicare SSI inpatient days divided by the total number of Medicare inpatient days plus the number of Medicaid, non-Medicare inpatient days divided by the total number of inpatient days (Centers for Medicare  Medicaid Services 2022). We accessed this measure directly from the CMS website.[100] Hospitals for which neither of these measures were available were dropped. Using these two definitions, we then constructed variables indicating whether a trial site (drawn from ClinicalTrials.gov) is a safety net hospital. To align with prior literature, for each measurement, hospitals within the top quartile of the state in which they are located were defined as a safety net hospital. Data from this section are used to produce Table VII and Appendix Table C16.

[98] Historical MRNA stock data can be downloaded here: https://finance.yahoo.com/quote/MRNA/history?p=MRNA

[99] AHA data were downloaded via the Wharton Research Data Service, under a license to Stanford University.

[100] Data available at https://data.cms.gov/provider-compliance/cost-report/hospital-provider-cost-report