

NBER WORKING PAPER SERIES

REPRESENTATION AND EXTRAPOLATION:  
EVIDENCE FROM CLINICAL TRIALS

Marcella Alsan  
Maya Durvasula  
Harsh Gupta  
Joshua Schwartzstein  
Heidi L. Williams

Working Paper 30575  
<http://www.nber.org/papers/w30575>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
October 2022, Revised July 2023

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, Jon Skinner, Stefanie Stantcheva, Adi Sunderam, Ebonya Washington, Crystal Yang 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, Boston University, ICSSI 2022, EAAMO 2022, and the U.S. Food and Drug Administration, for helpful comments. Nick Shankar, Lukas Leister, Anne Fogarty, Emma Ronzetti, Xingyou Ye, AnneMarie Bryson, 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. We note that Marcella Alsan was part of the Expert Committee for the National Academies of Science, Engineering and Medicine Report on Underrepresentation in Clinical Research referenced herein. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Marcella Alsan, Maya Durvasula, Harsh Gupta, Joshua Schwartzstein, and Heidi L. Williams. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Representation and Extrapolation: Evidence from Clinical Trials

Marcella Alsan, Maya Durvasula, Harsh Gupta, Joshua Schwartzstein, and Heidi L. Williams

NBER Working Paper No. 30575

October 2022, Revised July 2023

JEL No. D91,H5,I14

### **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 and evidence suggest 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.

Marcella Alsan  
Kennedy School of Government  
Harvard University  
79 John F. Kennedy St.  
Rubenstein Bldg R403  
Cambridge, MA 02138  
and NBER  
marcella\_alsan@hks.harvard.edu

Maya Durvasula  
Stanford University  
579 Jane Stanford Way  
Stanford, CA 94305  
mdurvas@stanford.edu

Harsh Gupta  
Stanford University  
hgupta13@stanford.edu

Joshua Schwartzstein  
Baker Library 461  
Harvard Business School  
Soldiers Field  
Boston, MA 02163  
jschwartzstein@hbs.edu

Heidi L. Williams  
Department of Economics  
6106 Rockefeller Center, Office 328  
Dartmouth College  
Hanover, NH 03755  
and NBER  
heidi.lie.williams@dartmouth.edu

An online appendix is available at <http://www.nber.org/data-appendix/w30575>

*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 disparities in both the production of clinical evidence and the eventual diffusion of products is commonplace (Ding and Glied 2022; Elhussein et al. 2022a,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 (Panel (a)) and are similarly underrepresented in prescriptions for newly approved medications (Panel (b)). Population is the implied benchmark in Figure I and we note Black patients are often even more underrepresented relative to their disease burden (Green et al. 2022). Although other groups have also been historically underrepresented, we focus on Black Americans for several reasons, including the history of racial discrimination and associated distrust, persistent racial disparities in health outcomes, and continued underrepresentation in research.<sup>1</sup>

[Figure I here]

While 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 would lead to higher drug demand.

---

<sup>1</sup>Female enrollment in clinical trials has been increasing over time and is currently comparable to female population share (see Appendix Figure B3), though gaps in certain conditions remain (Gupta 2022; Sosinsky et al. 2022; Feldman et al. 2019; Steinberg et al. 2021).

It also 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. A typical question from EBM training is: “Are the participants in the study similar enough to my patient?” (Masic, Miokovic and Muhamedagic 2008, p.222). 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 they 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.

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. To introduce 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.<sup>2</sup> 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 to realize any health gains. 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 a real clinical trial at two separate 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

---

<sup>2</sup>We informed physicians that the drugs were hypothetical so they would not try to prescribe them after the experiment.

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.

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. We also find in a separate but similar survey experiment that Black patients exposed to a representative trial were more likely to indicate that they want to participate in future clinical trials, and that they viewed the researchers as more trustworthy. This suggests that increasing representation might be one tool to help address medical mistrust. In contrast, and consistent with the model's prediction of diminishing returns to representation, we do not find significant effects associated with the trial composition for White patients. The combination of physician and patient results suggest that doctors are broadly acting as agents for their patients.

Survey experiments are important tools for uncovering peoples' mental models and perceptions (Stantcheva 2022a,b), but are also 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 that our aim was broadly to understand views on medical research, mirroring language from a non-profit dedicated to the same whenever feasible.<sup>3</sup> 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 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. Indeed the order of questions and salience of race might have played a role in the magnitudes of our effects. 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. Physicians educated at accredited medical colleges in the U.S. are explicitly taught to consider the applicability of trial findings to their own patients through EBM training (Blanco et al. 2014).

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 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 Academies of Science Engineering and Medicine report (NASEM 2022) and another comes from the association between more representative clinical

---

<sup>3</sup>Only 11.5 percent of physicians and 7.1 percent of patients attrited after consent and this was not differential across arms.

trials and higher prescribing rates for new drugs among Black patients (see Section VI.2 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 countries). One question is whether this hesitancy to extrapolate, especially among doctors, is a mistake. Given the current state of the literature and data availability, this does not seem to be a clear mistake. Manski, Mullahy and Venkataramani (2022) show, under mild assumptions regarding doctors’ objective function, that the inclusion of any predictive factor in clinical-decision making is welfare enhancing. Is race then in fact predictive of treatment effects? First, precisely because representation is so low, clinical trials offer limited direct evidence on this question. Green et al. (2022) review 290 new drug approvals in the FDA Drug Trials Snapshots data and 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 side effects and benefits, respectively. Ramamoorthy et al. (2015) report a higher rate of heterogeneous effects in a review of post-marketing analyses, finding such effects for nearly 20% of all new drugs. Second, because the medication mechanisms of action (*i.e.*, a drug’s pharmacodynamics) are often incompletely specified and evolving, 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 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 prescription gaps. 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. 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.

These findings then also imply that a firm could increase sales by recruiting a more representative sample. However, the trade-off in doing so is cost – our framework and evidence suggest 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, representation of Black patients then remains low – perpetuating doubt about whether trial findings

extrapolate to them and generating a cycle of underrepresentation.

Although policies that break this cycle of underrepresentation may take many forms, we discuss case studies of successful investments in what we call inclusive infrastructure. 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 with substantial federal investment, 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.

Our work contributes to a growing literature that seeks to understand the role of innovation in creating or exacerbating 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 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.

We also contribute to a literature on race and trust. People from different backgrounds may have different experiences (*i.e.*, different data to readily extrapolate from), and these experiences can lead to increased or decreased levels of trust that a variety of institutions work for them. Indeed, previous research has shown that differential beliefs in the returns to investment opportunities (Boerma and Karabarbounis 2023) contribute substantially to the persistence of the racial wealth gap (Derenoncourt et al. 2022). Research also indicates that historical exploitation, violence, and discrimination have led to distrust in the medical system and medical research (Alsan and Wanamaker 2018; Eli, Logan and Miloucheva 2019), declines in home ownership (Albright et al. 2021), and reduced participation in political processes (Williams 2022). Our paper provides a way to think about the consequences of these different experiences for trust more broadly, as the cycle of underrepresentation result applies to any process that includes a participation decision.

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 (36 percent), an order of magnitude more frequent than U.S. federal agencies (3 percent).<sup>4</sup> The remainder of clinical trials are sponsored by academic institutions, hospitals, and non-profit organizations.<sup>5</sup>

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.<sup>6</sup> 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

---

<sup>4</sup>See Ehrhardt, Appel and Meinert (2015) for evidence of the relative importance of industry sponsorship. Our estimates of the composition of clinical trials are drawn from ClinicalTrials.gov. We collected data on trials that both study products approved for sale in the United States and were subject to regulation by U.S. agencies. See Data Appendix H.1.1 for details.

<sup>5</sup>These institutions are flagged as “Other” in ClinicalTrials.gov. We reviewed institutions in this set to confirm that our interpretation of “Other” was correct.

<sup>6</sup>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.

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).<sup>7</sup> 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). 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 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. Obtaining proprietary information of these costs is difficult; however, several published studies and our own qualitative interviews with stakeholders provide corroborating evidence that White patients tend to require fewer resources and are thus much lower cost to recruit (see Appendix Section A.2 for details). Efforts to reach out to non-White communities typically involve additional staff, tailored recruitment materials, and new relationships with healthcare networks – all of which contribute to a comparatively high cost per enrollee (Marquez et al. 2003).

Two additional pieces of evidence provide some quantitative information on the size of these cost differences. First, consider the case of Moderna, which ran one of the highest-stakes clinical trials in recent history for its first-generation SARS-CoV-2 vaccine. In September 2020, the company announced that 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 illustration is the cost of recruiting experimental subjects for online 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 (Sections III and VI).

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

The cost differences described above may play a role in explaining the trial enrollment patterns observed in Figure I. Black patients make up just five percent of trial enrollees in the median clinical trial – far less

---

<sup>7</sup>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 the academic literature. For example, DiMasi, Grabowski and Hansen (2016) estimated the median cost of a Phase III trial as \$200 million.

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 benefits them, or to enroll in clinical trials (Table I).<sup>8</sup> 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).

[Table I here]

#### **II.1.4 Clinical Trials Data**

Upon successful completion of the three phases of clinical trials, sponsors submit new drug applications (NDA) 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 to patients with the targeted condition. Most trials are therefore powered to detect a mean difference in the primary endpoint 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.<sup>9</sup>

#### **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: 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). In particular, the U.S. currently lacks the price controls that other countries use to curtail spending and is permissive with respect to marketing. Given these features of the market, we focus on demand in the U.S., among physicians and patients.

---

<sup>8</sup>Note that these gaps are relatively constant when we control for income, education, and political affiliation (see Appendix Table C1). We also note, however, that conditioning on many characteristics may not always be appropriate when quantifying racial gaps (see Appendix Section A.1).

<sup>9</sup>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 U.S., 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.

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

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

Randomized controlled trials 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).<sup>10</sup>

After physicians complete their formal training, trial information is often accessed via multiple sources. These sources include ClinicalTrials.gov, which as of April 2019 received more than 215 million page views per month and 145,000 unique visitors daily. (See their website for additional details.) They also include academic journals, society or national practice guidelines, pharmaceutical representatives, medical conferences, and, more informally, online and in-person social networks. To maintain an active medical license, many primary care doctors participate in continuing medical education (CME). In addition to meeting requirements set by professional associations, doctors might wish to stay up to date with the literature for other reasons, including a desire to help their patients (Doximity 2014).

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

Patients learn about new drugs mainly through their physicians and via advertisements. The U.S. and New Zealand are the only two countries that allow firms to market medications directly to patients (Schwartz and Woloshin 2019). Between 2016 and 2018, firms spent \$17.8 billion on direct-to-consumer advertising (DTCA) associated with 553 unique drugs (U.S. Government Accountability Office 2021). Ads can be precisely targeted based on people’s search history and sometimes includes links to clinical information. 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 all major categories of disease.

Perhaps in part because of this outreach, data from Research!America show that 80 percent of Black respondents and 88 percent of White respondents had heard of clinical trials (Table I). Moreover, we document in our survey of primary care physicians that 72 percent report having ever 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.

---

<sup>10</sup>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 the 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).

### III Organizing Framework

The framework presented below 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. After presenting experimental tests of these predictions, we return to the framework in Section VI to try to understand why the underrepresentation of Black patients in clinical trials is so persistent.

#### III.1 Physicians and Patients

Physicians and patients use clinical trial information both to understand the benefits of a new treatment and to inform decisions about 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.<sup>11</sup>

##### III.1.1 Physician and Patient Beliefs

The assessments of patient-doctor dyad  $i$  are influenced by the current and historical trial data. Suppose the benefits to treatment for the patient in dyad  $i$  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:

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

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 (explicit or implicit) 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  $h$ :  $\hat{b}_i = \hat{b}(x_i; h)$ .

---

<sup>11</sup>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.

To focus 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  $t \in \{1, \dots, T\}$  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.<sup>12</sup> 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, we 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.<sup>13</sup> For doctors, this is consistent with evidence-based medicine (see 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  $x_i$  mattering.<sup>14</sup> We then have

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

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:

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

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

<sup>12</sup>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.

<sup>13</sup>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).

<sup>14</sup>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  $m$ . Incorporating such updating could strengthen the benefit of increasing Black representation.

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  $x_i$ .<sup>15</sup>

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

*Proof.* All proofs can be found in Appendix Section F.5. □

The intuition is straightforward: when a treatment works better than expected in the trial, people update their beliefs upwards on treatment efficacy.<sup>16</sup> 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.

We assume posteriors from the most similar previous treatment become the prior for a novel drug.<sup>17</sup>

---

<sup>15</sup>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. Specifically, it rules out physician- or patient-assumed heterogeneous trial efficacy as the mechanism driving our predictions. Relaxing this assumption would increase the importance of representation in our model.

<sup>16</sup>We focus on situations where the average trial efficacy exceeds prior-belief ratios for several reasons. First, it matches the focus on successful trials in our surveys. Second, doctors are asked to consider “favorable risk-benefit ratios” when recommending trials to their patients (Emanuel, Wendler and Grady 2000). Third, given the treatment approval process, patients tend to only have access to treatments that performed well in clinical trials. Fourth, it matches the empirical reality that trial results are typically only made public when successful (Turner et al. 2008, 2022; Driessen et al. 2015).

<sup>17</sup>Similar treatments could, for example, refer to treatments in the same category (drug class), or potentially all treatments for the same disease. 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).

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 (see Appendix Section F.4 for a numerical example). 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 differentiable cumulative distribution function  $F_\varepsilon(\cdot)$ .

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.<sup>18</sup>

Appendix Section F.3 formalizes two additional results on how beliefs impact behavior. First, Corollary F.2 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.<sup>19</sup> Second, Corollary F.3 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. Table II summarizes our theoretical predictions and how they connect to our empirical results, which we turn to next.

[Table II here]

## IV Experimental Design

### IV.1 Experimental Design

To test predictions from our theoretical framework, we conducted survey experiments – one with a sample of primary care physicians, and one with a sample of patients.<sup>20</sup> The experiments differed in important ways reflective of the different subject pools. Physicians, who are familiar with the task of evaluating new medications as part of standard practice, were asked to rate several hypothetical drugs. In each drug profile, the racial composition of the trial and efficacy were cross-randomized.<sup>21</sup> 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, a simpler exercise was presented: 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.<sup>22</sup> We

---

<sup>18</sup>As with Proposition 1, Corollary 1 restricts attention to how patients update given *successful* trials. While many trials fail, results from failed trials are less likely to be made public than results from successful trials (see footnote 16). In principle, patients could infer that “no news is bad news”. In practice, however, evidence suggests that people often do not make this type of inference even in simple laboratory settings (e.g., Jin, Luca and Martin 2021). We further note that there are additional mechanisms beyond those we formalize where more representative trials could increase Black patients’ willingness to participate in future trials, even if those trials are less successful than prior beliefs.

<sup>19</sup>The last result on diminishing returns requires mild regularity conditions on  $F_\epsilon(\cdot)$ .

<sup>20</sup>Appendix Figure B9 depicts the flow of the physician and patient surveys.

<sup>21</sup>We used hypothetical drugs instead of real drugs since there were not nearly enough real-world trials to include experimentally 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. In a complementary study, Ostrom (2022) reports that clinical trials funded by pharmaceutical companies report higher efficacy than when the same drug is used by a different study sponsor.

<sup>22</sup>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.

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 (*i.e.*, excluding hospitalists), 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.<sup>23</sup> 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).

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.<sup>24</sup>

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). There are no established guidelines that encourage different prescribing by race or ethnicity for patients with diabetes (Golden et al. 2012). However, there is a debate (as with other conditions) about the role genetic ancestry plays in its incidence (Parcha et al. 2022).

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

---

<sup>23</sup>We determine zip code rank using 5-year zip code-level population estimates reported in the 2019 American Community Survey.

<sup>24</sup>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.

exact duplicate) and drug names were selected from 15 alternatives.<sup>25</sup> 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 (e.g., metformin typically reduces A1c by 1-2 percentage points) (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).<sup>26</sup> Note that only efficacy and percent Black varied across the profiles, with all else held fixed.<sup>27</sup> 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.<sup>28</sup>

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.<sup>29</sup> 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. We also 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. Both campaigns were run by a non-profit, the Center for Information and Study on Clinical Research Participation (CISCRP).

## IV.1.2 Patient Survey Experiments

Patients were recruited from Lucid, an online survey platform frequently used in social science research and marketing (see the Data Appendix for more information on this platform). 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 included: (1) self-reported non-Hispanic

---

<sup>25</sup>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).

<sup>26</sup>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.

<sup>27</sup>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 and patient panel characteristics.

<sup>28</sup>Statistics on breakdown by sex were not provided in the drug profile. Although sex is an important characteristic, the policy issue of underrepresentation of women in trials is not as acute (see Appendix Figure B3).

<sup>29</sup>See Appendix Exhibits E4 and E5 for the exact question wording shown to physicians and a link to the survey.

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.<sup>30</sup> 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). For the experiment assessing the new medication, we introduced consequentiality by explicitly encouraging patients to answer truthfully, and noting that their responses would be used to generate a personalized report they could download and share with their primary care provider. Approximately 42 percent of patient respondents downloaded the personalized report.

We began the experimental module by providing basic details about the clinical trial process. Before randomization, we informed respondents that new medications to treat blood pressure are frequently studied by researchers. We noted these new therapies typically aim to improve blood pressure control, reduce complexity, or decrease side effects 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 shown 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 criterion – 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.<sup>31</sup> 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, in text and graphic form, 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.<sup>32</sup> 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 describe the reasons for this limited confidence.

---

<sup>30</sup>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.

<sup>31</sup>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.

<sup>32</sup>The exact question wording shown to patients and a link to the survey can be found in Appendix Exhibits E7 and E8.

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

Our survey experiment on clinical trial participation followed the above design but occurred several months later, using a separate group of patients. In this second study, the outcome of interest was a respondent's stated willingness to participate in a new trial that was similar to the one they had been shown. After respondents provided this information, we asked multiple-choice questions designed to elicit views on the financial or medical consequences of trial participation, on whether the trial would produce new or relevant knowledge, on data privacy, and on researcher trustworthiness.

## V Experimental Data and Results

### V.1 Sample Characteristics

We invited 12,192 physicians to participate in the study.<sup>33</sup> 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 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.<sup>34</sup>

We recruited 275 patients diagnosed with hypertension to provide views on a novel treatment: 139 Black and 136 White respondents. Respondents are comparable to individuals with hypertension in the Medical Expenditure Panel Survey (MEPS); in Appendix Table C5, we document that Black and White respondents in our survey are broadly similar to MEPS respondents by age, geography, income, and insurance status, although there were relatively more female respondents. Black respondents had slightly higher levels of college education and White respondents were less educated than in the MEPS data (Blewett et al. 2019).<sup>35</sup> We recruited another 272 participants to the clinical trials participation experiment. There was no significant imbalance or differential attrition across arms for any survey (see Appendix Tables C2, C7, C8, C9, and C10).

---

<sup>33</sup>In total, 4.7 percent of emails bounced and 1.8 percent of those invited started the survey. 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 (Richardson, Dominowska and Ragno 2007; Kanich et al. 2009).

<sup>34</sup>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).

<sup>35</sup>Our main results are robust to including person weights derived from a nationally representative survey, the Medical Expenditure Panel Survey (Appendix Table C6).

## 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:

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

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 ( $\rho_k$ ), mechanism of action 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.

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

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

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 oneself) 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.<sup>36</sup> 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

---

<sup>36</sup>Non-standardized outcomes and continuous updating outcomes yield similar results, which are gathered in the Appendix Table C12. 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.

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 who 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 enter linearly. Figure II explores these relationships nonparametrically by interacting quartiles of patient percent Black with the treatment and plotting the total effect (main effect plus 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. 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 B10 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 panel percent Black, nor between the other demographic categories. Appendix Figure B11 additionally demonstrates few associations between physician-specific responses to representative trials and their own background characteristics.

[Table III here]

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, 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).

Lastly, we turn to our main results from the follow-up experiment, which investigated the relationship between beliefs about trial representation and willingness to participate in future clinical trials. Results are reported in Column (1), Panels (a) and (b) of Appendix Table C11. We find that exposure to the treatment – data on a more representative trial – increases Black patients’ stated willingness to participate in similar future blood pressure studies by 0.385 standard deviation units. There was no significant effect for White patients and the difference in treatment effects across the two groups was significant ( $p$ -value = 0.038). We discuss potential mechanisms for these results in Section V.2.3.

[Figure II here]

## V.2.2 Representation and Disparities

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. For Black hypertensive patients, beliefs about how much the drug will lower blood pressure are within 1 mm of the range of the reported clinical effect for 33 percent of respondents, compared to almost 60 percent of White hypertensive respondents. 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 substantially on trial results, reporting a perceived effectiveness for their own health that is similar to the study finding. In contrast, Black patients are more willing to accept that reported efficacy under study conditions captures the drug’s effectiveness for their own health when the sample is more representative (K-S test  $p$ -value 0.026).

[Figure III here]

[Figure IV here]

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 intentions 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 here]

### **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 of “offshoring” clinical trials (Petryna 2009).

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 that a drug tested on one sample would not work equally well in a different context. We provided a set of multiple choice responses that allowed respondents to indicate concerns about biological factors, socioeconomic and environmental factors, or trust in the trial. Participants were also allowed to select “other” and asked to provide open-text answers.

Results are reported in Table IV. Panel (a) presents views from Black patients and doctors who treat them regarding extrapolation 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: ranging from 7.0 percent among PBP to 15.4 percent among PWP. 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 (Panel b) is slightly higher than their counterparts in Panel (a). Third, when providing a rationale for why a drug might work differently across samples, a nontrivial share selected 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 the 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.”

Returning to the experimental results, we find that Black patients who view others as trustworthy were significantly more likely to want to ask their doctor about the new medication (Appendix Table C13 Column (3)). In addition, we find that the representative treatment increases Black patients’ willingness to participate in future clinical trials, as well as their views on the trustworthiness of the trial researchers (Appendix Table C11, Panel (a) Column (2)). The same pattern does not hold for White patients (Panel (b) Column (2)).

[Table IV here]

## 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.*, sharing personalized reports with their doctors) experiments. The majority of physicians and nearly half of all patient respondents requested access to these reports, suggesting participants indeed valued them. For the patient survey, all respondents had been diagnosed with hypertension and thus had limited incentives to distort their 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 presented above (see Appendix Tables C15 and C6 for experimental results from physicians and patients, respectively). Appendix Tables C4 and C16 show that patients and doctors who downloaded or requested the report are statistically similar to other respondents.

The second key feature that reduces concerns about social desirability or experimenter demand effects is that we pre-specified heterogeneous effects by the patient’s race and the racial composition of the provider’s patient panel. If social desirability was playing a large role, patterns might be similar 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, information presented in our physician survey closely resembled the demographic information presented in biomedical publications and regulatory publications (*e.g.*, the FDA Drug Trial Snapshots database).

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.<sup>37</sup> Our follow-up donation survey finds that the amount physicians allocate to the enrollment campaign targeting underrepresented minorities is strongly and significantly associated with physician-specific coefficients on representation (Table V) 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.

[Table V here]

## V.2.5 Threats to External Validity

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

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 to our patient respondents immediately after the intervention about the rationale for their responses; sentiment analysis reported in Appendix Table C18 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 then 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

---

<sup>37</sup>We sent a follow-up survey to physicians after at least a week, 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).

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 information on trial representation matters in practice, we document that it affects prescribing intention and updating from trial results. In settings outside of our experiments, evidence that Black Americans are skeptical of research institutions and medical technologies – FDA approved and investigational – is widespread. We tabulate survey respondents consistent with these patterns in Table I. Qualitative comments from physicians in our study, as well as those drawn from a recent NASEM report, also suggest that representation plays a role in how doctors practice medicine (see Appendix Tables D1 and D2 and Appendix Figure B13).

## V.2.6 Robustness

We probe the robustness of our findings for physicians in Appendix Table C15. 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 either to physicians who answer our follow-up donation question or to those who request a copy of our report to NIH and NASEM (see Columns (4) and (5)). Column (6) shows that findings on representation are not sensitive to the addition of controls selected using double-selection LASSO linear regression (Chernozhukov et al. 2018). We also find that the order of profiles presented to physicians does not substantially impact how they respond to the treatment (Appendix Figure B14).

Additional results from our physician sample are presented in Appendix Table C17. 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.<sup>38</sup> 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 MEPS. Panel (c) indicates that our results are robust to including LASSO-selected controls.

---

<sup>38</sup>See Appendix Section A.3 for additional discussion.

## VI Discussion and Conclusion

The theoretical and experimental analysis sheds light on the potential benefits of increasing representation of Black patients in clinical trials to both patients and pharmaceutical companies. Given these benefits, why does such underrepresentation persist?

One hypothesis would be that this underrepresentation persists because of a combination of a relative lack of information and distrust of doctor recommendations between Black and White patients. However, the racial participation gap in clinical trials is much larger than would be implied by the observed gaps in trust and information reported in Table I.<sup>39</sup> This section uses a combination of theory and case studies to analyze why this gap is so persistently large, extending the earlier theoretical and experimental analysis to study the costs and benefits to firms conducting clinical trials. In the process, this section fleshes out a potentially important intertemporal externality associated with a history of underrepresentation.

### VI.1 Why Might Underrepresentation Persist?

Suppose pharmaceutical firms seek to maximize the expected profit from a given experimental drug trial and can choose their recruitment strategy (see Appendix Section F.2 for details). They have access to a status-quo technology for recruiting patients to clinical trials. Under this technology, a racial gap in perceived treatment benefits increases the racial gap in trial *participation* relative to the gap in trial *recruitment* (Proposition F.3). In other words, firms using the status quo technology anticipate a higher refusal rate from Black vs. White patients. Firms could choose to incur a fixed cost  $f > 0$  to increase Black representation from its level under the status quo by making investments that reduce the marginal costs of inviting more Black participants. We refer to these investments as building “inclusive infrastructure.” Our theoretical and empirical results suggest firms would see value from such investment: due to diminishing returns to representation, it could increase demand among Black patients and their doctors without significantly decreasing demand among White patients or their doctors. However, the returns to such investment may not be completely internalized by any given firm: it increases perceived benefits for all similar treatments in the future, including those developed by *other* firms.<sup>40</sup> The externalities a firm’s *current* recruitment decisions have on other firms’ *future* recruitment costs enables a cycle of underrepresentation.

**Proposition 2.** *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*

---

<sup>39</sup>Table I suggests that Black patients are about 90% as likely as Whites to have heard of a clinical trial and about 94% as likely to say they would enroll if a doctor recommended it. If they only hear about clinical trials when doctors point them out, then this implies they should participate in clinical trials at about 85% the rate of White patients. But Figure I implies that the share of Black patients in clinical trials relative to their population share is around 33% that of White patients.

<sup>40</sup>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. Such an externality suggests that firms may underinvest in such technology relative to what is socially optimal. See Model Appendix Sections F.2 and F.4 for details.

*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 is illustrated with a numerical example in Model Appendix Section F.4.

Together, the theoretical and empirical results (summarized in Table II) are suggestive of 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 increase trial representation actively. (4) Trials today are not representative for Black patients.<sup>41</sup> (5) And the cycle continues.

## VI.2 Case Studies

The theoretical analysis suggests that investments in inclusive infrastructure may help break such a cycle of underrepresentation. Here, we combine quantitative and qualitative evidence, including insights drawn from informal 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.<sup>42</sup> 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 prescriptions of new drugs to Black Americans across various conditions.

[Figure VI here]

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 research networks across the U.S. by the National Cancer Institute (NCI) and National Institute of Allergy and Infectious Diseases (NIAID), respectively.<sup>43</sup> 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).

---

<sup>41</sup>While Proposition 2 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.

<sup>42</sup>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 few pharmaceuticals intended to prevent/treat such deaths.

<sup>43</sup>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.

The history of these research networks—and their specific forms of investment—shed light on differences in contemporary outcomes across disease areas. Investment in cancer research has, historically, been driven by top-down investments into academic medical centers, including efforts in the “War on Cancer” that began with the National Cancer Act of 1971 (Mukherjee 2010). Beginning in 1972, motivated by a Howard University study documenting “an astounding increase in cancer mortality among the nation’s Black population in recent years,” the National Cancer Institute (NCI) invested in efforts to understand the burden of cancer mortality across racial groups (Henschke et al. 1973; Wailoo 2011). Following the passage of the 1993 NIH Revitalization Act, however, investigators receiving NCI funding reported struggling to comply with new rules regarding minority representation in clinical trials because NCI funding could not be used for “ancillary” study costs, including reimbursements for patient expenses, resources for advertising and outreach, and funding for patient navigators and counselors.

In contrast to the top-down development of federal cancer research infrastructure, research into HIV/AIDS has been shaped by community involvement and activism. Activists pushed researchers to alter standard protocols for research, calling for accelerated approvals and emergency access to medicine, introduction of surrogate endpoints that could proxy for other clinical markers, and greater emphasis on representation in trial recruitment (Epstein 1996). In parallel, political, religious, and community leaders worked to combat the stigma associated with links between HIV/AIDS and homosexuality, especially in Black communities, thus creating opportunities for individuals to seek access to experimental therapies (Robertson 2006; Royles 2020). At a 1990 community forum on clinical trials held in San Francisco, ACT UP / San Francisco member Michelle Roland called for a “revolution in clinical trial design,” in which activists and scientists designed “realistic clinical trials that do a better job of meeting people’s needs” (as recounted in Epstein (1996), Chapter 7). In response to demands from activists, the ACTG and the National Institute of Allergy and Infectious Diseases adopted the practice of seeking community involvement at each trial site when developing protocols, prioritizing long-term relationships outside of academic medical centers (Kagan et al. 2012).

Table VI substantiates these anecdotes and makes clear how site selection shapes trial composition. Amongst U.S.-based trial sites listed in the ClinicalTrials.gov database, sites that enroll for HIV/AIDS are approximately 11 (16) percentage points more likely to be located at a safety net hospital than sites that recruit for cancer (Alzheimer’s Disease Related Dementias (ADRD)). Unsurprisingly, the demographic characteristics of the trial sites also differ. Appendix Tables C20 and C21 report information on the demographics of HIV/AIDS, cancer, and ADRD research centers at the hospital service area level for all clinical trials and for specific networks.<sup>44</sup> Trial sites recruiting for cancer have, on average, a 10.5 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.<sup>45</sup>

[Table VI here]

---

<sup>44</sup>See Appendix Table VI for more comparisons.

<sup>45</sup>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. 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).

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 Division of AIDS (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 (CAB) (Strauss et al. 2001).<sup>46</sup> The CAB meets regularly with trial investigators and consults on proposed protocols. Our discussion with HVTN leadership suggests that DAIDS requirements have important spillover effects: although firms are not obligated to comply, industry sponsors often engage with communities to benefit from existing recruitment networks.

The stark differences in trial composition for cancer and HIV/AIDS highlight the extent to which active, large-scale investments in inclusive infrastructure, in addition to incentives, can be important for improving health disparities. Figure VI Panel (b) demonstrates a positive relationship between greater representation in trials and prescribing rates.<sup>47</sup> 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 potentially useful template for industry and regulators.

### VI.3 Concluding Comments

Motivated by persistent, substantial racial disparities in both clinical trial enrollment and prescriptions for new drugs, we investigated the consequences and causes of underrepresentation of Black patients in medical research. Consistent with a theoretical model of similarity-based extrapolation, 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 gaps close. The results suggest that a feedback loop exists between representation in a process and subsequent decision-making. Such a cycle of underrepresentation could apply more widely to any data-driven participation or take-up decision.

HARVARD KENNEDY SCHOOL AND NATIONAL BUREAU OF ECONOMIC RESEARCH, UNITED STATES

STANFORD UNIVERSITY, UNITED STATES

STANFORD UNIVERSITY, UNITED STATES

HARVARD BUSINESS SCHOOL, UNITED STATES

STANFORD UNIVERSITY AND NATIONAL BUREAU OF ECONOMIC RESEARCH, UNITED STATES

## Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

---

<sup>46</sup>Although some institutions maintain a CAB for cancer trials, the CAB requirement at DAIDS is unique (National Institute of Allergy and Infectious Diseases 2022).

<sup>47</sup>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 may in turn have pull incentives on innovation as per Acemoglu et al. (2006), Finkelstein (2004), and Acemoglu and Linn (2004).

## References

- Acemoglu, Daron, and Joshua Linn.** 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *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.
- Agency for Healthcare Research and Quality.** 2022. "MEPS Prescribed Medicine Files 1996-2019." [https://meps.ahrq.gov/mepsweb/data\\_stats/download\\_data\\_files.jsp](https://meps.ahrq.gov/mepsweb/data_stats/download_data_files.jsp), Accessed 9/05/2022.
- 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.
- Albright, Alex, Jeremy A. Cook, James J. Feigenbaum, Laura Kincaid, Jason Long, and Nathan Nunn.** 2021. "After the Burning The Economic Effects of the 1921 Tulsa Race Massacre." *NBER Working Paper No. 28985*, July.
- Alsan, Marcella, and Marianne Wanamaker.** 2018. "Tuskegee and the Health of Black Men." *Quarterly Journal of Economics*, 133(1): 407–455.
- American Diabetes Association.** 2020. "Section 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2021." *Diabetes Care*, 43: S111–S124.
- 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.
- Barnato, Amber E.** 2017. "Challenges In Understanding And Respecting Patients' Preferences." *Health Affairs*, 36(7): 1252–1257.
- 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>.
- Boerma, Job, and Loukas Karabarbounis.** 2023. "Reparations and Rersistent Racial Wealth Gaps." *NBER Macroeconomics Annual*, 37(1): 171–221.
- 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, 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.
- 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, Pragma 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.
- Daw, Nathaniel D.** 2014. “Advanced Reinforcement Learning.” *Neuroeconomics*, 299–320.
- Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick.** 2022. “Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860-2020.” Working paper.
- 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*.
- DiMasi, Joseph A., Henry G. Grabowski, and Ronald W. Hansen.** 2016. “Innovation in the Pharmaceutical Industry: New Estimates of R&D 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.
- Driessen, Ellen, Steven D. Hollon, Claudi L. H. Bockting, Pim Cuijpers, and Erick H. Turner.** 2015. “Does Publication Bias Inflate the Apparent Efficacy of Psychological Treatment for Major Depressive Disorder? A Systematic Review and Meta-Analysis of US National Institutes of Health-Funded Trials.” *PLoS One*, 10(9).

- 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.
- Elhoussein, 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.** 2022a. “Racial-Ethnic Disparities in Uptake of New Hepatitis C Drugs in Medicare.” *Lancet Regional Health - Americas*, 6(100111): 1147–1158.
- Elhoussein, 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.** 2022b. “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.
- Emanuel, Ezekiel, David Wendler, and Christine Grady.** 2000. “What Makes Clinical Research Ethical?” *JAMA*, 283(20): 2701–2711.
- Epstein, Steven.** 1996. *Impure Science: AIDS, Activism, and the Politics of Knowledge*. Berkeley, CA:University of California Press.
- 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.” *Quarterly Journal of Economics*, 119(2): 527–564.
- Foster, Andrew D, and Mark R Rosenzweig.** 2010. “Microeconomics of Technology Adoption.” *Annual Review of Economics*, 2(1): 395–424.
- Gilboa, Itzhak, and David Schmeidler.** 1995. “Case-Based Decision Theory.” *Quarterly Journal of Economics*, 110(3): 605–639.
- 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.
- Gupta, Harsh.** 2022. “Do Female Researchers Increase Female Enrollment in Clinical Trials?” *Working paper*.
- 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*.
- Henschke, U.K., L.D. Leffall Jr, C.H. Mason, A.W. Reinhold, R.L. Schneider, and J.E. White.** 1973. “Alarming Increase of the Cancer Mortality in the U.S. Black Population (1950-1967).” *Cancer*, 31(4): 763–768.
- 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.
- IQVIA.** 2015. “Global Medicines Use in 2020: Outlook and Implications.” IMS Institute for Healthcare Informatics Technical Report.
- 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.” *Quarterly Journal of Economics*, 125(2): 515–548.
- Jin, Ginger Zhe, Michael Luca, and Daniel Martin.** 2021. “Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure.” *American Economic Journal: Microeconomics*, 13(2): 141–173.
- 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.
- Kagan, Jonathan M., Scott R. Rosas, Rona L. Siskind, Russel D. Campbell, Daniel Gondwe, Daniel Munroe, William M.K. Trochim, and Jeffery T. Schouten.** 2012. “Community-researcher Partnerships at NIAID HIV/AIDS Clinical Trials Sites: Insights for Evaluation and Enhancement.” *Progress in Community Health Partnerships: Research, Education, and Action*, 6(3): 311–320.
- 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.
- 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.
- 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.
- Malmendier, Ulrike, and Laura Veldkamp.** 2022. “Information Resonance.” *Working paper*.
- Manski, Charles F., John Mullahy, and Atheendar Venkataramani.** 2022. “Using Measures of Race to Make Clinical Predictions: Decision Making, Patient Health, and Fairness.” *NBER Working Paper No. 30700*.
- Marquez, Miriam A., Joan M Muhs, Ann Tosomeen, B Lawrence Riggs, and L Joseph Melton III.** 2003. “Costs and Strategies in Minority Recruitment for Osteoporosis Research.” *JBMR*, 18(1): 3–8.
- 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 R&D: Missing Innovation from Regulating Female Enrollment in Clinical Trials.” *Working paper*.
- 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.” *Quarterly Journal of Economics*, 117(3): 735–774.
- Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer.** 2008. “Coarse Thinking and Persuasion.” *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 Institute of Allergy and Infectious Diseases.** 2022. “DAIDS Community Engagement.” <https://www.niaid.nih.gov/daids-ctu/community-engagement-NEW>, Accessed 10/11/2022.
- National Institutes of Health.** 2017. “NIH’s Definition of a Clinical Trial.” <https://grants.nih.gov/policy/clinical-trials/definition.htm>, Accessed: 10/10/2022.
- Oostrom, Tamar.** 2022. “Funding of Clinical Trials and Reported Drug Efficacy.” *Electronic Health Economics Colloquium*.
- 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.
- Papageorge, Nicholas W.** 2016. “Why Medical Innovation is Valuable: Health, Human Capital and the Labor Market.” *Quantitative Economics*, 7(3): 671–725.
- Parcha, Vibhu, Brittain Heindl, Rajat Kalra, Adam Bress, Shreya Rao, Ambarish Pandey, Barbara Gower, Marguerite R. Irvin, Merry-Lynn N. McDonald, Peng Li, Garima Arora, and Pankaj Arora.** 2022. “Genetic European Ancestry and Incident Diabetes in Black Individuals: Insights From the SPRINT Trial.” *Circulation: Genomic and Precision Medicine*.
- Petryna, Adriana.** 2009. *When Experiments Travel*. Princeton, NJ:Princeton University Press.
- 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.
- Ramamoorthy, A, M.A. Pacanowski, J Bull, and L Zhang.** 2015. “Racial/Ethnic Differences in Drug Disposition and Response: Review of Recently Approved Drugs.” *Clinical Pharmacology and Therapeutics*, 97(3): 263–273.

- 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.
- Robertson, Ronald.** 2006. "Civilization." *Theory, Culture, and Society*, 23(2).
- 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.
- Royles, Dan.** 2020. *To Make the Wounded Whole: The African American Struggle Against HIV/AIDS*. Chapel Hill, NC: The University of North Carolina Press.
- 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 Berling, 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.
- 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.** 2022a. "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible." *NBER Working Paper No. 30527*, (30527).
- Stantcheva, Stefanie.** 2022b. "Understanding Tax Policy: How do People Reason." *Quarterly Journal of Economics*, 136(4): 2309–2369.
- STAT News.** 2020. "Representation and Diversity in Clinical Trials." <https://www.statnews.com/representation-diversity-clinical-trials/>, Accessed: 10/11/2022.
- 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. Goepfinger, 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.” <https://www.cnn.com/2020/09/04/moderna-slows-coronavirus-vaccine-trial-t-to-ensure-minority-representation-ceo-says.html>.
- Turner, Eric H., Andrea Cipriani, Toshi A. Furukawa, Georgia Salanti, and Ymkje Anna de Vries.** 2022. “Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy: Updated Comparisons and Meta-analyses of Newer Versus Older Trials.” *PLoS Medicine*, 19(1).
- Turner, Eric H., Annette M. Matthews, Eftihia Linardatos, Robert A. Tell, and Robert Rosenthal.** 2008. “Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy.” *New England Journal of Medicine*, 358: 252–260.
- 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.
- Wailoo, Keith A.** 2011. *How Cancer Crossed the Color Line*. New York, NY:Oxford University Press.
- 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 Bickdeli, 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.
- Williams, Jhacova.** 2022. “Historical Lynchings and the Contemporary Voting Behaviors of Blacks.” *American Economic Journal Applied Economics*, 14(3): 224–253.

## Tables

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 (0.963)	3.082 (0.822)	-0.253***
Heard of Clinical Trial	0.796 (0.374)	0.875 (0.339)	-0.079***
Would Enroll in Clinical Trial if Doctor Recommends	0.783 (0.384)	0.837 (0.379)	-0.054***
Trust Not Reason for Lack of Enrollment	0.432 (0.463)	0.536 (0.514)	-0.104***
Science is Beneficial	0.284 (0.419)	0.383 (0.493)	-0.099***
Would Get FDA-Approved Vaccine	2.907 (1.024)	3.069 (1.099)	-0.163

*Notes:* Table reports the survey responses from Black and White U.S.-based respondents for a set of questions regarding science. Data are from a national survey conducted by the non-profit Research!America over 2013, 2017 and 2021. *Heard of Clinical Trial*, *Trust*, and *Science is Beneficial* are dichotomous variables. Other variables are on an ordinal scale. See Data Appendix for details on variable construction. Standard deviations are in parentheses. \*, \*\*, \*\*\* refer to statistical significance at the 10, 5, and 1 percent level, respectively.

Table II: Summary of Theoretical Predictions and Empirical Results

Theory	Predictions	Exhibits	Result Summary
Prop. 1.1; Cor. F.2.1	Perceived benefits and demand for a new medication are increasing in trial-reported efficacy.	<b>Table III</b>	A 1 sd increase in efficacy increases physician prescribing intention by 0.28 sd.
Prop. 1.2; Cor. F.2.2	Perceived benefits and demand for a new medication are increasing in representation of similar patients in clinical trials.	<b>Table III</b>	<ul style="list-style-type: none"> <li>• For physicians, a 1 sd increase in representation increases prescribing intention by 0.11 sd.</li> <li>• 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”) by 0.78 sd and 19.9 pp, respectively.</li> </ul>
Prop. 1.3; Cor. F.2.3	Diminishing returns to representation.	<b>Figure II(d); Table III</b>	<ul style="list-style-type: none"> <li>• 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.</li> <li>• 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.</li> </ul>
Cor. F.3	<ul style="list-style-type: none"> <li>• There are White-Black gaps in perceived benefits and demand for a new medication.</li> <li>• Increasing Black representation in clinical trials narrows these gaps.</li> </ul>	<b>Figure III; Figure IV; Figure V</b>	<ul style="list-style-type: none"> <li>• 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.</li> <li>• 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 1 pp less likely to load on the signal than White patients and this difference is statistically indistinguishable from 0.</li> </ul>
Cor. 1	Groups that were historically underrepresented in successful trials have a lower propensity to participate in trials today than historically well-represented groups.	<b>Appendix Table C11; Section VI.2</b>	<ul style="list-style-type: none"> <li>• For Black patients, being assigned to the representative treatment increases their stated willingness to participate in similar future blood pressure studies by 0.39 sd.</li> <li>• Historically, HIV/AIDS trials were more representative than cancer trials. Recent HIV/AIDS trials are associated with a higher percent Black representation than recent cancer trials.</li> </ul>

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

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

<b>Panel A: Primary Care Physicians</b>						
	<u>Relevance</u>	<u>Prescribing</u>	<u>Relevance</u>	<u>Prescribing</u>	<u>Relevance</u>	<u>Prescribing</u>
	No Controls		Main Specification		Share Black Interactions	
	(1)	(2)	(3)	(4)	(5)	(6)
Representation	0.163*** (0.039)	0.179*** (0.036)	0.109*** (0.029)	0.107*** (0.029)	0.007 (0.038)	-0.005 (0.039)
Efficacy	0.165*** (0.038)	0.229*** (0.039)	0.189*** (0.029)	0.281*** (0.032)	0.179*** (0.036)	0.285*** (0.043)
Representation × Patient Percent Black					0.004*** (0.001)	0.004*** (0.001)
Efficacy × Patient Percent Black					0.000 (0.001)	-0.000 (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
<b>Panel B: Patients</b>						
	<u>Relevance</u>		<u>Ask Doctor</u>		<u>Loading on Signal</u>	
	Black Patients	White Patients	Black Patients	White Patients	Black Patients	White Patients
	(1)	(2)	(3)	(4)	(5)	(6)
Representative Treatment	0.781*** (0.167)	0.172 (0.159)	0.021 (0.077)	0.006 (0.079)	0.199** (0.083)	-0.057 (0.086)
$p$ -value: Black Patients=White Patients		0.008***		0.893		0.030**
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* and *Efficacy* with the reported percent of patients that are Black in the physician's panel and the main effect is included but not reported. 137 physicians participated in the experiment each assessing eight oral antidiabetic 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 1. *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.

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

<b>Panel A: Black Patients and Their Physicians (PBP)</b>						
<b>White to Black Patients</b>	<u>Confidence</u>				<u>Rationale</u>	
	Not at All (1)	Some (2)	Moderate (3)	High (4)	Perceived Biol. Factors (5)	Perceived Social & Envir. Factors (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%
<b>Panel B: White Patients and Their Physicians (PWP)</b>						
<b>Offshored to U.S. Patients</b>	<u>Confidence</u>				<u>Rationale</u>	
	Not at All (1)	Some (2)	Moderate (3)	High (4)	Perceived Biol. Factors (5)	Perceived Social & Envir. Factors (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 C14 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.

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

	(1)	(2)
Coefficient on Representation	1.279*** (0.449)	1.229*** (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.

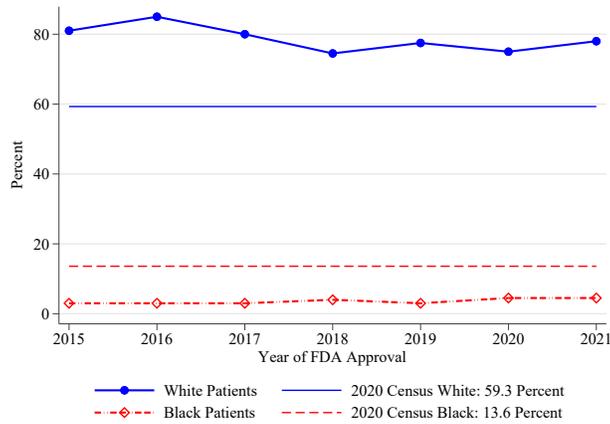
Table VI: Trial Sites and Safety Net Hospitals

	DSH Index		UCMP Care	
	(1)	(2)	(3)	(4)
HIV/AIDS (Cancer Comparison)	0.110*** (0.008)		0.019*** (0.007)	
HIV/AIDS (ADRD Comparison)		0.161*** (0.012)		0.054*** (0.010)
Constant	0.475	0.423	0.176	0.141
Observations	197,240	6,804	182,929	5,997

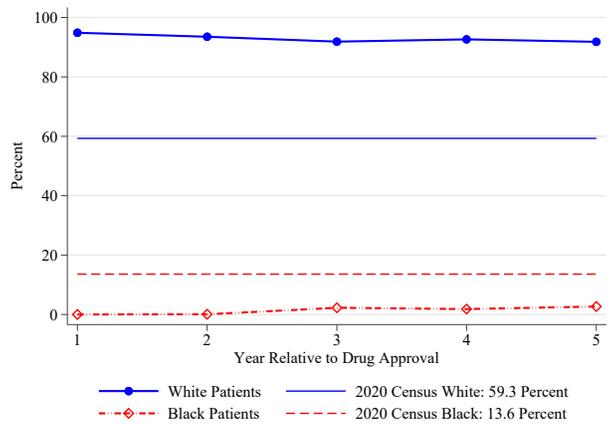
*Notes:* Table reports OLS estimates from a regression of an indicator for whether a trial site is located at a safety net hospital (SNH). Each observation represents a specific site associated with a unique clinical trial and the data are limited to Cancer, HIV/AIDS, and ADRD trials. Following Popescu et al. (2019), we define a SNH as a hospital in the state's top quartile of Medicaid and Medicare Supplemental Security Income inpatient days historically used to determine Medicare Disproportionate Share Hospital (DSH) payments (Columns (1) and (2)); and uncompensated (UCMP) care costs (as a percentage of total operating expenses) (Columns (3) and (4)). See the Data Appendix for more detailed definitions of these variables. *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 Alzheimer's Disease and Related Dementias (ADRD). See Appendix Table C19 for a *Cancer (ADRD Comparison)*. Trial site information is drawn from ClinicalTrials.gov. See Data Appendix H.1.1 and H.3.8 for details. Robust standard errors are in parentheses. \*, \*\*, \*\*\* refer to statistical significance at the 10, 5, and 1 percent level, respectively.

# Figures

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



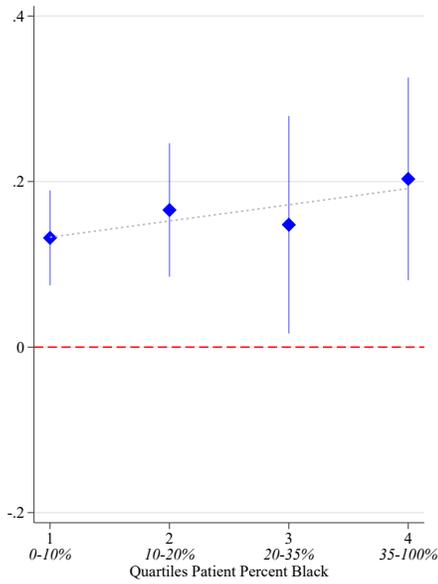
(a) Clinical Trials Participation



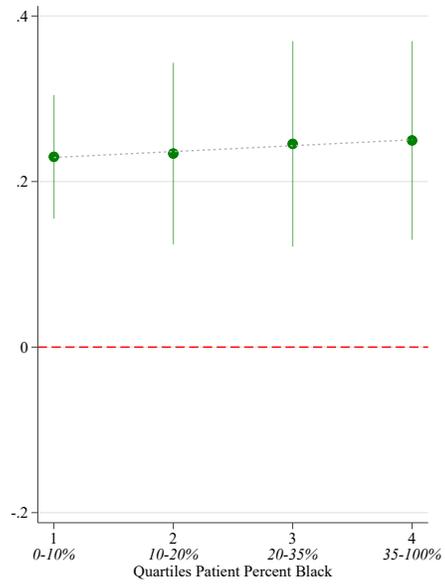
(b) Prescriptions of New Drugs

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 median new drug prescription percentage by race in each year relative to its approval. Straight lines in Panels (a) and (b) 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)). Panel (a) is drawn from the FDA Drug Trials Snapshots data, and Panel (b) and is from the Medical Expenditure Panel Survey data (Agency for Healthcare Research and Quality 2022). Appendix Figure B1 plots (a) using a longer time series from ClinicalTrials.gov. Appendix Figure B2 plots the distribution of race in trials using both the ClinicalTrials.gov and FDA Drug Trials Snapshots data sets. Appendix Figure B4 plots prescribing rates of new drugs per 1000 individuals in each racial group.

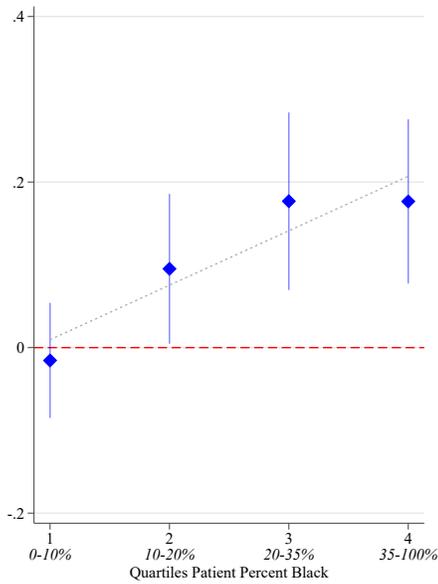
Figure II: Heterogeneity Among Physicians by Racial Composition of Patient Panel



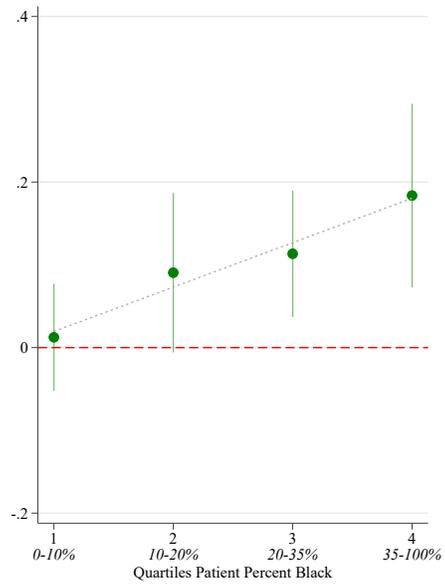
(a) Efficacy on Relevance



(b) Efficacy on Prescribing Intention



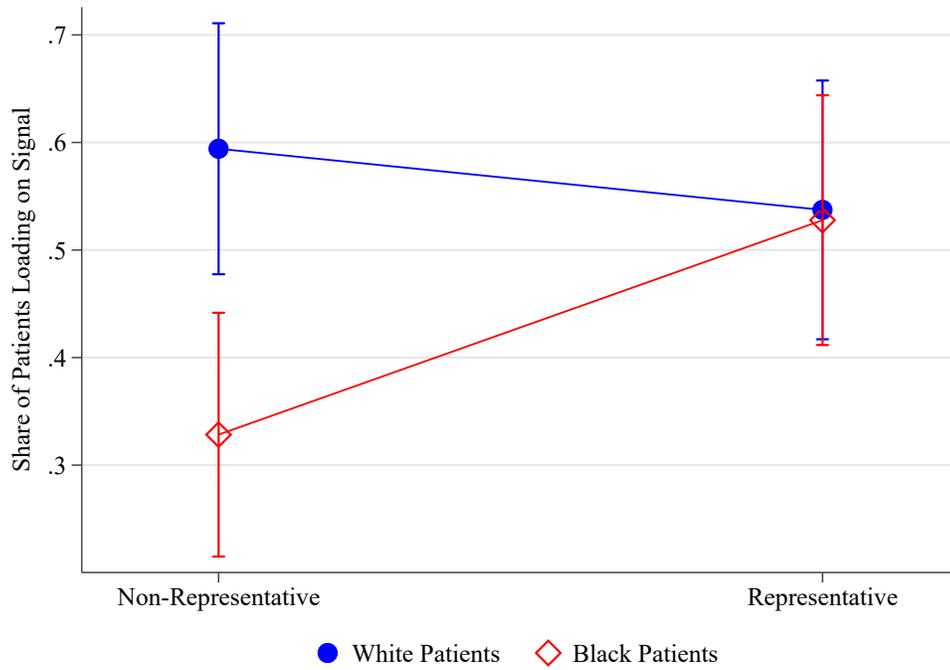
(c) Representation on Relevance



(d) Representation on Prescribing Intention

Notes: Figure plots OLS estimates for two outcomes – *Relevance* (Panels (a) and (c)) and *Prescribing Intention* (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.

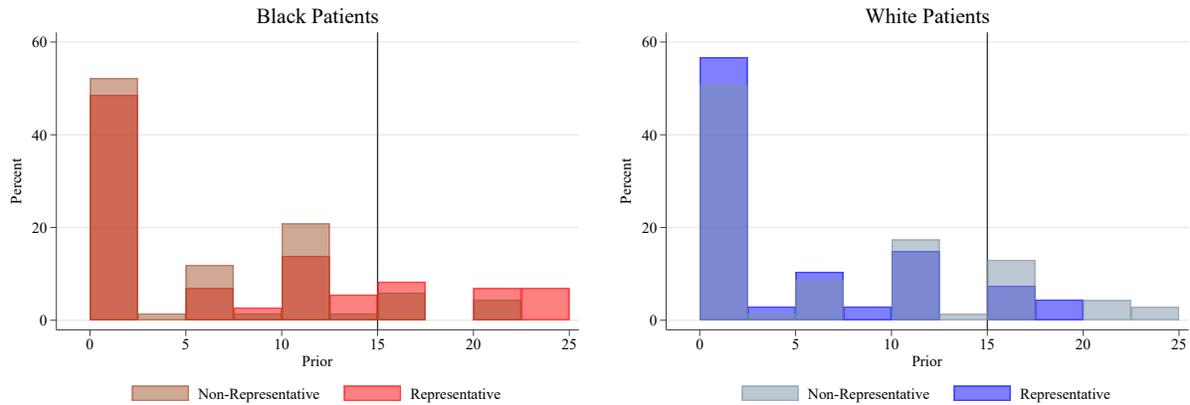
Figure III: Loading on Signal by Race and Treatment Status



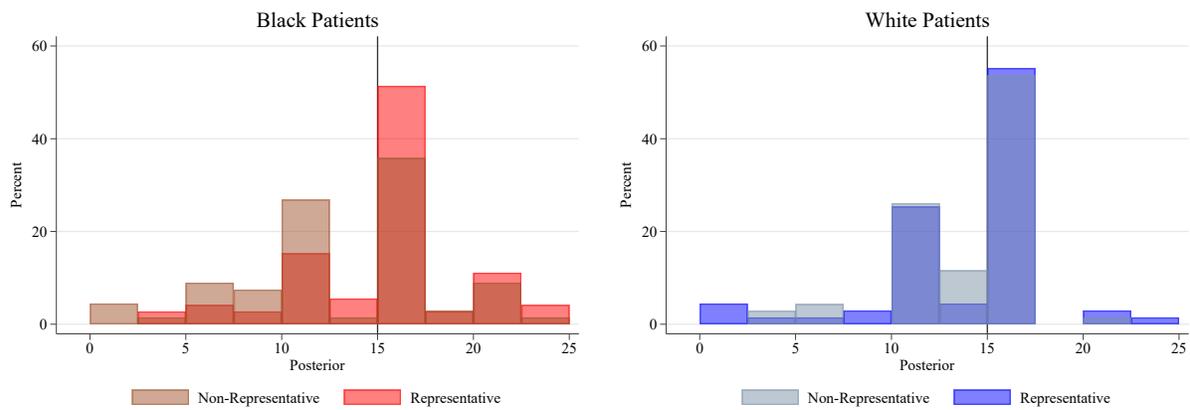
Notes: Figure plots the share of respondents who “Load on Signal” – whose posteriors are within 1 mmHg of the reported drug efficacy in our intervention (15 mmHg) – by race and treatment group. *Load on Signal* is an indicator variable that takes a value of one if the respondent’s posterior was between 14 and 16, and zero otherwise. The x-axis reports values for two groups of respondents: non-representative trials with <1 percent Black patients and representative trials with 15 percent Black patients. Results are plotted separately by respondent race. 95 percent confidence intervals are included.

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

*Panel A. Prior Beliefs*

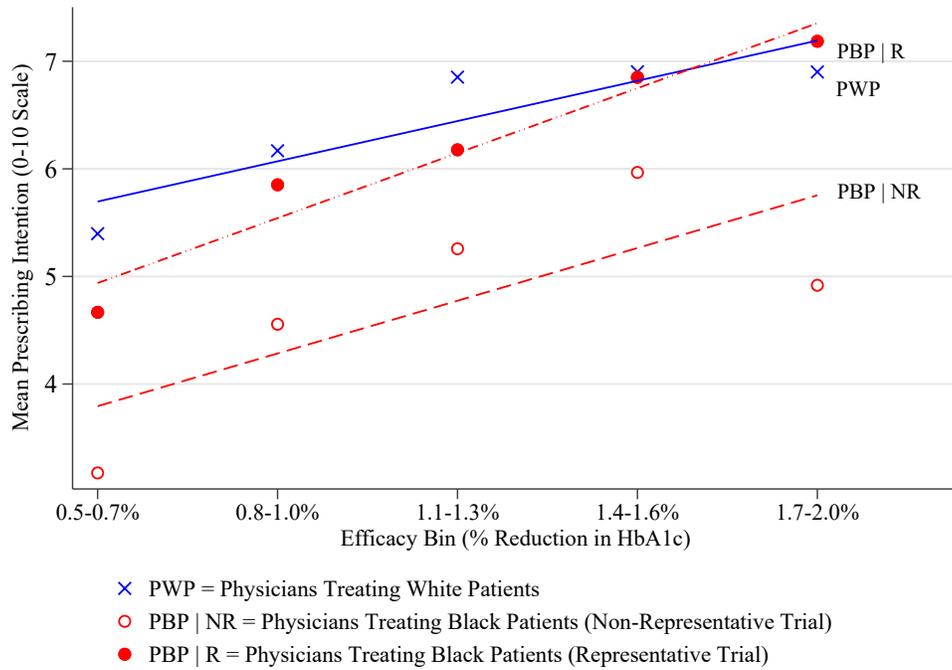


*Panel B. Posterior Beliefs*



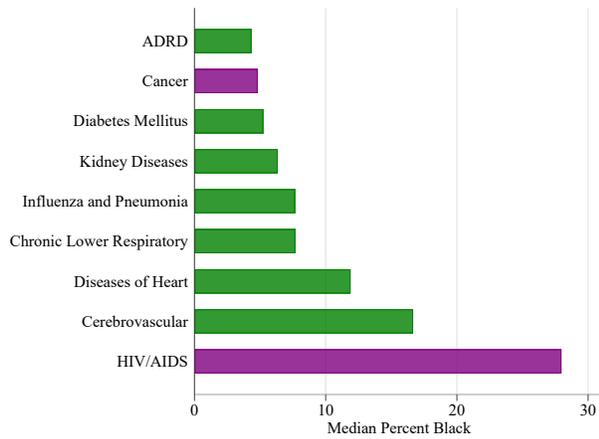
*Notes:* Figure plots the prior and posterior distribution of beliefs about the perceived efficacy of the new antihypertensive medication for the patient’s own condition by respondent’s race and assigned treatment status (trial shown is either non-representative or representative). The signal on efficacy 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, a Kolmogorov-Smirnov test rejects the null that the posteriors are identical across arms ( $p$ -value = 0.026). For White patients, a Kolmogorov-Smirnov test fails to reject the null that the posteriors are identical across arms ( $p$ -value=0.789).

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

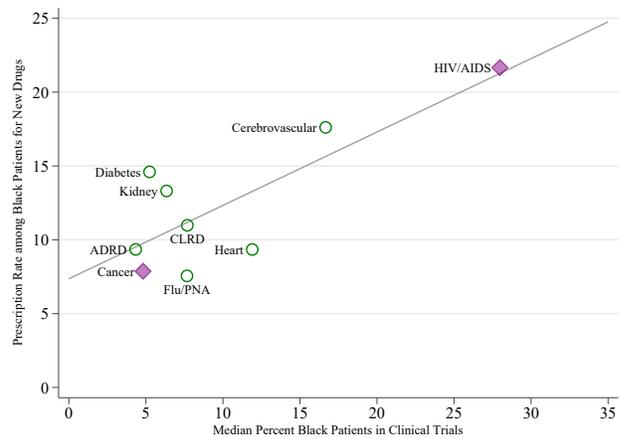


Notes: Figure plots the relationship between *Efficacy* and *Prescribing Intention* (on a 0-10 scale) by patient composition and percent Black of trial subjects in the profiles shown to physicians. *PBP* (Physicians treating Black Patients) denotes physicians who report above the median percent Black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

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



(a) Median Percent Black Patients 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 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 includes outpatient cancer supportive therapies. CLRD, Diabetes, Heart, Kidney, and Flu/PNA indicate Chronic Lower Respiratory Diseases, Diabetes Mellitus, Diseases of Heart, Kidney Diseases, and Influenza and Pneumonia, respectively. Prescription data are from the Medical Expenditure Panel Survey. Observations associated with cancer and HIV/AIDS are denoted with diamonds (purple). See Data Appendix for details.