Journal of Health Economics 100 (2025) 102968

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhe

Investigator racial diversity and clinical trial participation☆

Marcella Alsan [a,*], Romaine A. Campbell [b], Lukas Leister [c], Ayotomiwa Ojo [d]

[a] Kennedy School of Government, Harvard University, , Cambridge, MA 02138, United States of America
[b] Brooks School of Public Policy, Cornell University, Ithaca, NY 14850, United States of America
[c] Department of Economics and Business, Universitat Pompeu Fabra, , 08005 Barcelona, Spain
[d] Northwestern Memorial Hospital, Chicago, IL 60611, United States of America

ARTICLE INFO

Keywords:
Diversity
Equity
Clinical trials
Innovation

ABSTRACT

We investigate whether increased racial diversity of clinical trial principal investigators could increase the enrollment of Black patients, which currently lags population and disease-burden. We conducted a survey experiment in which respondents were shown a photo of a current NIH investigator in which race (Black/White) was randomized. Sex was also randomized as a relevant benchmark. Black respondents reported 0.35 standard deviation units higher interest in participating in a clinical study led by a race concordant investigator (a 12.6% increase). Sex concordance had no effect. Further analyses indicate that perceived trustworthiness and attractiveness are the most important factors explaining these results.

1. Introduction

Despite comprising 13 percent of the U.S. population and suffering disproportionately from premature morbidity and mortality, Black Americans make up only 5 percent of clinical trial participants (Alegria et al., 2021). In qualitative research, Black Americans commonly cite a lack of trust in academic and research institutions as well as investigators as the most significant attitudinal barrier to their participation in research (Scharff et al., 2010). Low enrollment of racial and ethnic minorities may compromise the generalizability of research findings and affect the opportunity for minority groups to benefit from medical innovation (Manski et al., 2023). While rigorous causal evidence on the implications of low representation for patient health is sparse, simulations and immunologic evidence suggest that lower clinical trial participation among racial minorities has contributed to growing racial disparities in diabetes morbidity and cancer survival (Basu and Gujral, 2020; Awidi and Al Hadidi, 2021). In addition, a recent randomized online experiment among physicians and patients found that a lack of representativeness in clinical trials negatively affects Black patients' beliefs about stated drug efficacy and reduces the willingness of physicians who treat Black patients to prescribe new medications (Alsan et al., 2024).

Strengthening diversity in clinical trials has been a policy priority for leading scientific institutions and policymakers. In 2015, the U.S. Food and Drug Administration (FDA) launched a five-year plan to increase the racial diversity and representativeness of clinical trials (FDA, 2014). However, a 2022 analysis of clinical trial data found that Black patients continue to be underrepresented, with less than 20 percent of trials reporting treatment or side effects data for Black patients (Green et al., 2022). In response, the FDA

☆ 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. Nikhil Shankar and Anne Fogarty provided excellent research assistance. The study is approved by IRB at Harvard University (IRB21-0864) and registered at the AEA RCT Registry (AEARCTR-0011926) and ClinicalTrials.gov (NCT04938895). We gratefully acknowledge financial support from the National Science Foundation, United States under Grant No. DGE1745303.

* Corresponding author.
E-mail addresses: marcella_alsan@hks.harvard.edu (M. Alsan), rcampbell@cornell.edu (R.A. Campbell), lukas.leister@upf.edu (L. Leister), ayotomiwa.ojo@nm.org (A. Ojo).

https://doi.org/10.1016/j.jhealeco.2025.102968
Received 16 August 2024; Received in revised form 9 January 2025; Accepted 9 January 2025
Available online 31 January 2025

has issued draft guidance (FDA, 2022), and the U.S. Congress has introduced legislative proposals to improve diversity in clinical trial research through initiatives such as public awareness campaigns and the use of mobile technologies (H.R.6000, H.R.5030, H.R.6584, H.R.3085).

Effective strategies to inform and encourage Black patients to participate in trials are needed to complement these policy measures. The recent *National Academies of Science Engineering and Medicine Report* on "Improving Representation in Clinical Trials and Research" outlined several potential strategies to accomplish this aim. These recommendations included increasing remuneration for study participation and fast-tracking patent applications for products tested on diverse samples. The report also highlighted the importance of ensuring a diverse and inclusive workforce, particularly in leadership roles, for all those involved in clinical trials research (National Academies of Sciences, Engineering, and Medicine, 2022, p. 133). Notably, both the National Institute on Minority Health and Health Disparities and the National Institute on Drug Abuse Office of Diversity and Health Disparities have established initiatives to recruit, support, and fund minority students and researchers (NIMHD, 2022; NIH, 2022a). However, the effectiveness of improved racial diversity among principal investigators for increasing clinical trial participation among Black adults has not been established empirically.

This project aims to fill the gap. We conducted a randomized survey experiment on a representative sample of Black Americans, randomizing exposure to race concordant photographs of current National Institutes of Health (NIH) investigators and assessing stated willingness to participate in a future clinical trial. Sex concordance was cross-randomized as a benchmark because it constitutes another salient dimension of similarity between subjects and investigators. We demonstrate that exposure to a race concordant investigator increases willingness to participate in medical research among Black adults by 0.35 standard deviation units or 12.6 percent. In contrast, sex concordance has no statistically significant effect.

Our results are robust to the inclusion of respondent demographics and machine learning (ML)-generated ratings of image characteristics, such as image quality and the emotions expressed by the investigator (Serengil and Ozpinar, 2021; Ocampo, 2022). We conduct several additional checks to validate our results. A common concern with survey experiments is that findings may be driven by demand effects. To assess this, we exclude the 6 percent of participants who mentioned race- or sex-related terms in an open-text question about the study's purpose from our analysis, and our results persist. We rerun our main specification excluding each investigator in turn and demonstrate that the resulting estimates cluster around our baseline estimate and also show that estimated investigator fixed effects within treatment category are statistically indistinguishable from each other, suggesting that our findings are not driven by outlier investigators. Finally, we show that the racial concordance effect is robust to alternative definitions of the outcome variable as well as the inclusion of controls for sex concordance with the investigator, having a medical condition, risk aversion, trust in others, and altruism.

So, what *is* driving our effect? To explore mechanisms, we query subjects on the perceived attributes of the investigator on several margins after introducing the experimental variation. These include perceived age, attractiveness, educational attainment, relative quality, and trustworthiness. We chose these attributes based on prior literature supporting their importance in persuasion. Bertrand et al. (2010) randomized features of a direct mailer for loans in South Africa and found that an attractive photo was equivalent to a 25 percent lower interest rate on the loan. Educational attainment and quality can signal professionalism or legitimacy and have experimentally been shown to garner a return in the medical labor market (Deming et al., 2016). Age might be perceived as greater experience or outmoded information and thus may be a positive or negative signal. Lastly, trustworthiness is important given the history of exploitation and abuse particularly targeting Black Americans (Brandt, 1987; Roberts, 1997; Alsan and Wanamaker, 2018; Washington, 2006). We decompose our overall effect by iteratively including each investigator feature in our main specification individually. We find that the perceived trustworthiness of the researcher has a substantial explanatory effect, almost doubling the fit of the model (*i.e.*, R-squared from 0.116 to 0.218) and decreasing the race concordant treatment effect by about one-third. Perceived attractiveness displays a similar pattern, although the two are only modestly correlated in the data. Perceived quality matters much less than either trustworthiness or attractiveness, and no other attribute is shown to matter. A formal decomposition method following Gelbach (2016) confirms the importance of trustworthiness and attractiveness.

Our paper is related to several strands of literature, including the importance of trust in the delivery of medical care (Alsan and Wanamaker, 2018; Banerjee et al., 2020) and the role of racial concordance in improving health outcomes (Alsan et al., 2019; Alsan and Eichmeyer, 2024; Anderson et al., 2020; Frakes and Gruber, 2022; Hill et al., 2020; Greenwood et al., 2020). We add to this literature by rigorously examining the decision to participate in medical research, which is important for innovation and confidence in research results (Alsan et al., 2024). Moreover, to our knowledge, we provide the first experimental evidence that sex concordance does not seem to affect patients' perceptions in a medical context — a finding consistent with existing observational studies (Lau et al., 2021; Takeshita et al., 2020; Street et al., 2008). Methodologically, we borrow from Ludwig and Mullainathan (2023) by combining real-world images with ML-rated attributes and using the output in a choice experiment for a policy-relevant topic.

The paper proceeds as follows. In Section 2, we provide more information on racial gaps in investigator funding and leadership as well as previous studies on concordance. In Section 3, we describe our sample and intervention, including investigator image selection and outcomes. We then present results in Section 4. Finally, we conclude with a discussion of the potential implications of our results.

2. Background

Several studies have highlighted the gap between the demographics of clinical trial samples and those in the population who either have the relevant condition or would be primary targets for prevention. In a recent review of cancer trials, Nazha et al. (2019)

find that only 5 percent of the samples were Black across trials for novel checkpoint inhibitors. While COVID-19 vaccine trial samples were more representative at nearly 9 percent Black (Artiga et al., 2022), these trials were unique in terms of the considerable scrutiny from the press, pressure by the U.S. government, and major advance purchase agreements. Recent trials for Alzheimer's, a disease that has a larger impact on Black communities, continue to display the pre-COVID pattern for trial representativeness: Eli Lilly's 1,736-person donanemab trial included only 19 Black participants who received the drug (Reardon, 2023).

Against this backdrop, diversity among investigators has been championed as a potential path to reduce the racial recruitment gap. Diverse investigators might ask different research questions, develop more successful outreach, enrollment, and retention strategies for minorities in trials, and increase trust both in the specific studies they are conducting and in the research process more broadly. Yet, there is no causal evidence to establish the effectiveness of this approach. Previous research has documented that physician-patient race, sex, or ethnic concordance can improve patient satisfaction (Takeshita et al., 2020; Cooper et al., 2003), patient trust in health information (Loeb et al., 2023), utilization of preventative services (Alsan et al., 2019), mortality among those with manageable illnesses (Frakes and Gruber, 2022), neonatal outcomes (Greenwood et al., 2020), and survival of heart attacks (Greenwood et al., 2018). Religious concordance between medical experts and patients increased intent to take up the COVID-19 vaccine among American Christians (Chu et al., 2021). Because the investigator-trialist relationship is distinct from the doctor-patient relationship, evidence specific to this setting is needed.

Yet the topic is difficult to study for two main reasons. First, investigators are not randomly assigned to studies. Second, systemic discrimination and structural inequality have led to persistent racial disparities in access to education, promotion, and funding opportunities (Laurencin et al., 2023; National Academies of Sciences, Engineering, and Medicine, 2023). Consequently, Black investigators comprise less than 4 percent of NIH-funded researchers (NIH, 2022b). The funding gap was first documented by Ginther et al. (2011) who found that African American applicants were 13 percentage points less likely to receive NIH investigator-initiated funding than White investigators. While recent updates suggest that racial representation and funding rate gaps have narrowed since 2011, non-White NIH principal investigators remain significantly underrepresented and underfunded (NIH, 2023). Therefore, we devised an online trial in which we randomized profiles of real Black and White NIH investigators to subjects.

3. Experimental design

The goal of the experiment was to test whether racial concordance with clinical trial investigators affects patient willingness to participate in clinical studies. In addition, we sought to probe the mechanisms underlying a potential effect. We randomly assigned subjects to an image of an NIH investigator and elicited their willingness to participate in a future study as well as their perceptions of the principal investigator. This section describes our experimental design choices in detail.

3.1. Study flow, description of treatment, and outcome measurement

Study Flow: The overall study flow is presented in Appendix Figure B1. After providing informed consent, subjects answered screening questions (described in Section 3.2). Eligible respondents were then randomized to the treatment. Following the intervention, we queried subjects on their willingness to participate in a trial led by the randomly assigned investigator — our primary outcome of interest. To assess potential mechanisms, we subsequently asked subjects to rate the investigator on various characteristics that prior literature suggests are important for persuasion (details below). Finally, we asked about the respondent's current access to medical care, risk preferences, and included open-response questions about the study's purpose.

Treatment Description: The intervention is shown in Fig. 1. The race (Black/White) and sex (Male/Female) in the image of a current NIH investigator were cross-randomized. The investigator's head shot was centered and an accurate description of their role as a lead investigator of the NIH was provided.

Selection of Investigator Images: Professional head shots of NIH intramural research program (IRP) investigators were sourced from pictures publicly available on the IRP website (NIH, 2022c). The IRP has over 12,000 principal investigators conducting research and is the world's largest biomedical research institution. The investigators are employed by the NIH directly. As Black investigators are less represented on the website and comprise less than 4 percent of NIH-funded researchers, we found fewer usable pictures for Black investigators than for White investigators (NIH, 2022b). We ultimately selected five pictures of Black males, seven of Black females, ten of White males, and eight of White females for the experiment.[1]

To ensure the comparability and quality of investigator images, we followed a principled image selection process. First, we hand-selected pictures of investigators that were taken in front of a neutral background and that featured a neutral expression. Second, we used ML-generated ratings of the investigator photos to confirm that picture quality, investigator age, and the emotions conveyed by the NIH investigators were similar across demographic groups. The ML-generated investigator emotions were Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. We used the Python packages deepface (Serengil and Ozpinar, 2021) and image-quality (Ocampo, 2022) to generate these ratings. The image file size was used as an additional proxy for quality. ML-generated ratings of the investigator pictures were balanced across race (Appendix Table A1). Female investigators were generally rated more favorably (*i.e.*, rated as younger, happier, and less neutral) than their male counterparts (Appendix Table A2). As the same investigator images were shown to both male and female respondents, each image was part of both the sex concordant treatments, so that balance across treatments was assured.

[1] Images used in the experiment are available to interested parties by specific request.

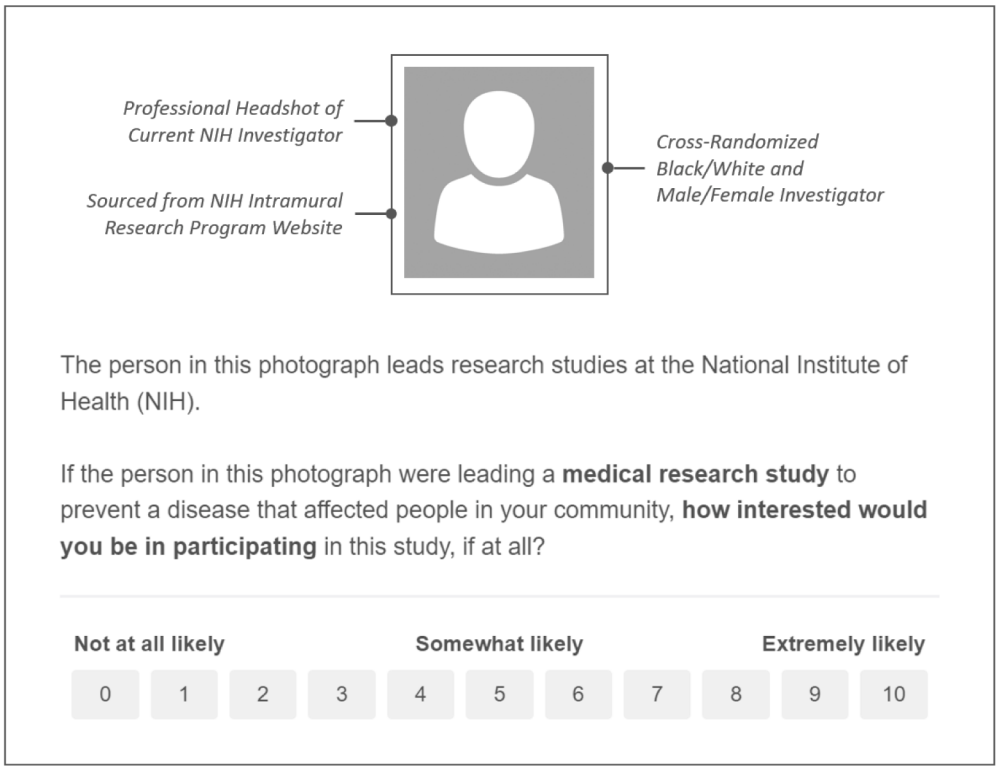

Fig. 1. NIH investigator randomization. *Notes:* Depiction of the Randomized photo of an NIH investigator and the question relating to the primary outcome.

Outcome Description: After viewing the image, participants were asked, "If the person in this photograph were leading a medical research study to prevent a disease that affected people in your community, how interested would you be in participating in this study, if at all?" We chose this generic language so that we could elicit responses from all subjects, not just those with a specific condition. Our primary outcome variable is the answer to this question measured on a ten-point Likert scale and standardized for analysis (see Appendix Table C1). We present the distribution of interest in clinical trial participation in Appendix Figure B2.

To probe mechanisms, we asked subjects to rate investigator age, education, trustworthiness, attractiveness, and relative quality. Investigator age was shown on an interval scale (25–34 years, 35–44 years 45–54 years, 55–64 years, 65+ years), investigator education was measured by highest educational attainment (options ranged from less than a high school degree to completion of a professional degree) and attractiveness, quality, and trustworthiness were measured on five-point Likert scales. See Appendix Table C1 for additional details on the precise wording of questions, answer options, and coding for analysis. For our analysis, we standardized all perceived investigator characteristics except education. Education was coded as an indicator variable for whether the respondent selected graduate degree (*i.e.*, M..D., J.D., or Ph.D.), though results are not sensitive to this recoding.

Demand Effects: The experiment was specifically designed to mitigate experimenter demand effects. As the respondents were only shown one investigator, it would have been difficult for them to discern the rationale of the study. Indeed, only 5.6 percent of respondents included race- or sex-related terms in response to the open-ended question "What do you think this study was about?". We present a wordcloud of subject responses in Appendix Figure B3 with the dominant term being "health". Nevertheless, we exclude such respondents in robustness checks (see Section 4.2).

3.2. Recruitment

We recruited Black survey participants using CloudResearch in February 2022. CloudResearch is an online survey vendor frequently used in social science research. The platform relies on an aggregation of opt-in market research panels coupled with additional data quality checks and selection criteria (Chandler et al., 2019). We limited the sample to Black subjects given their under-representation in clinical trials and previous research showing the importance of concordance in Black populations (Alsan et al., 2021; Alsan and Eichmeyer, 2024; Loeb et al., 2023). Upon completing the consent process, participants answered a set of questions to determine eligibility based on nationality (U.S.-born), self-identified race (non-Hispanic Black), and age (25–64). We also screened out respondents who failed our attention check. Approximately 27 percent of individuals who began the survey were screened out due to eligibility criteria or demographic quotas. Eligible respondents went on to answer basic demographic questions covering sex, income, education, state of birth, health insurance, employment, and existing medical conditions. Our pre-specified recruitment target of 300 participants was determined by power calculations based on responses from a pilot survey.

Table 1
The Effect of NIH Investigator Race or Sex Concordance on Clinical Trial Participation Interest.

Panel A: Racial Concordance				
	(1)	(2)	(3)	(4)
Racial Concordance	0.293***	0.294***	0.322**	0.350***
	(0.110)	(0.109)	(0.134)	(0.132)
Control Mean	−0.147	−0.147	−0.147	−0.147
Panel B: Sex Concordance				
	(1)	(2)	(3)	(4)
Sex Concordance	−0.013	−0.037	−0.000	−0.010
	(0.112)	(0.111)	(0.111)	(0.108)
Control Mean	0.006	0.006	0.006	0.006
Observations	323	323	323	323
Respondent Covariates	No	Yes	No	Yes
Image Covariates	No	No	Yes	Yes

Notes: OLS estimates of the willingness to participate in a future clinical trial led by race (Panel A) or sex concordant (Panel B) NIH investigators. Outcomes are standardized to have a mean of 0 and standard deviation of 1. Column (1) includes only an indicator for treatment. Column (2) adds respondent characteristics (age, squared age, education, employment, and South/non-South region). Column (3) adds ML-rated image characteristics (predicted investigator emotion, age, image quality, image size), and Column (4) includes both. All specifications include an intercept that is omitted in the table. Robust standard errors are shown in parentheses.

3.3. Descriptive statistics, balance, and attrition

Our experimental sample exceeded our target of 300 and consisted of 323 Black individuals who fulfilled our eligibility criteria (see Section 3.2), passed our quality check, and completed the survey. Summary statistics are presented in Appendix Table A3. Our sample is broadly representative of the U.S. non-Hispanic Black population by age, sex, geography, education level, employment, and health insurance status (Appendix Table A3).[2] Respondents were, on average, 44 years old and about 40 percent had a high school degree or less education. The only characteristic that differed significantly from the U.S. population was employment status, with our survey participants being less likely to hold formal employment. Respondent characteristics were balanced across race and sex concordant/discordant arms (see Appendix Tables A4 and A5). Attrition after randomization was low — among respondents who were exposed to the treatment (i.e., saw an image of an investigator), less than one percent did not complete the survey. In Appendix Table A6, we show that attrition was balanced across treatment arms.

4. Estimation and results

4.1. Main findings

Given the cross-randomized experimental design, we estimate the following equation:

$$Y_{si} = \beta_0 + \beta_j Concordance^{j}_{si} + X_s'\gamma + M_i'\delta + \epsilon_{si} \quad (1)$$

where s is the subject, i is the investigator image, and j signifies either racial or sex concordance. We first estimate separate equations for racial and sex concordance. In subsequent specifications, we simply control for respondent sex as we fail to find differences by sex concordance. Y_i is the standardized outcome capturing interest in participating in a clinical trial. X_s is a vector of respondent covariates including age, squared age, education level, employment status, and an indicator for whether the respondent resides in the South. M_i is a vector of the ML-generated investigator image characteristics, such as predicted investigator emotion, age, and image quality, as well as image size.

Table 1 displays our main results. Panel (A) reports findings for racial concordance and Panel (B) reports results for sex concordance. Column (1) includes no controls except an indicator for treatment status. We find that survey respondents randomized to an image of a racially concordant NIH investigator are 0.293 standard deviation units more likely to participate in a clinical trial led by the investigator compared with a standardized mean of −0.147 among respondents assigned to a racially discordant investigator. The same pattern does not hold for Panel (B). We do not find detectable effects of sex concordance on participation. Moving across the columns, we add controls for respondent and ML-rated image characteristics. The final column includes both sets of characteristics and is our preferred specification. Race concordant NIH investigators increase willingness to participate by 0.350 standard deviation units. Model fit increases from an R-squared of 0.022 to an R-squared of 0.116. Random assignment to a sex concordant NIH investigator does not significantly affect interest in study participation in any specification.

[2] We used respondents to the 2019 American Community Survey as a comparison and limited to non-Hispanic Black Americans with internet access and English proficiency which were requirements for the participation in our survey.

Table 2
Racial and Sex Concordance on Perceived of Investigator Characteristics.

Panel A: Racial Concordance					
	Age	Education	Quality	Attractiveness	Trustworthiness
	(1)	(2)	(3)	(4)	(5)
Racial Concordance	−0.365***	0.131**	0.495***	0.408***	0.309**
	(0.124)	(0.066)	(0.134)	(0.139)	(0.138)
Control Mean	0.061	0.410	−0.163	−0.175	−0.132
Panel B: Sex Concordance					
	(1)	(2)	(3)	(4)	(5)
Sex Concordance	0.150	−0.076	−0.044	−0.122	−0.073
	(0.107)	(0.055)	(0.112)	(0.111)	(0.113)
Control Mean	−0.049	0.494	0.022	0.050	0.031
Observations	323	323	323	323	323
Respondent Covariates	Yes	Yes	Yes	Yes	Yes
Image Covariates	Yes	Yes	Yes	Yes	Yes

Notes: OLS estimates of ratings of investigator characteristics by race (Panel A) or sex (Panel B) concordance. In Column (2), the outcome is an indicator for completion of a graduate degree. Outcomes in Column (1) and Columns (3)-(5) are standardized to have a mean of 0 and standard deviation of 1. All specifications include respondent characteristics (age, squared age, sex, education, employment, and South/non-South region) and ML-rated image characteristics (predicted investigator emotion, age, image quality, image size). All specifications include an intercept that is omitted in the table. Robust standard errors are in parentheses.

Next, we explore potential underlying mechanisms for the effect of investigator racial concordance. To do so, we use Eq. (1), placing investigator characteristics on the left-hand side (Table 2). On average, subjects perceived race concordant investigators as younger and rated them more positively on all characteristics, including 0.495 standard deviation units more qualified, 0.408 standard deviation units more attractive, and 0.309 standard deviation units more trustworthy (Panel A). Sex concordance does not significantly affect perceived image characteristics (Panel B).

We then investigate whether perceived investigator characteristics partially account for our estimated racial concordance effect. In Fig. 2, we plot the coefficient estimates from our main specification (hereafter, Model 0) and alternative specifications. Recall that Model 0 regresses clinical trial participation on racial concordance as well as respondent and ML-rated investigator image characteristics (Table 1, Column (4)). In Models 1–5, we iterate through respondents' ratings of perceived investigator characteristics, adding trustworthiness, attractiveness, quality, education, and age to regression models individually (Panel A).[3] In Model 6, we saturate the model with *all* perceived investigator characteristics (Panel B). In Panel (C), we plot the R-squared, demonstrating how model fit varies with the inclusion of different perceived investigator characteristics.

We find that perceived investigator trustworthiness is important for unpacking the racial concordance effect. The inclusion of trustworthiness reduces the racial concordance effect by 0.103 standard deviation units and increases the fit of the model from an R-squared of 0.116 to an R-squared of 0.218 (Fig. 2, Model 0 and Model 1). We can reject the null hypothesis that the coefficients are the same in the two models (p-value = 0.026). Similarly, the inclusion of either attractiveness or quality causes a decrease of similar magnitude but less significance in the racial concordance effect, reducing the estimated coefficient by roughly 0.110 and 0.093 standard deviation units, respectively (Fig. 2, Model 2 and Model 3). Again, we can reject the null hypothesis that the coefficients are the same across the models (Model 0 vs. Model 2: p-value = .005 and Model 0 vs. Model 3: p-value = 0.016). They also both increase model fit but to a lesser degree than trustworthiness (Fig. 2, Panel C). Perceived investigator education and age have a negligible impact.

In a multivariate regression of clinical trial participation on treatment and *all* perceived investigator characteristics (Model 6), the R-squared rises to 0.244 (Panel C). In this specification, racial concordance increases interest in future trial participation by 0.213 standard deviation units, compared to 0.350 standard deviation units in Model 0. Among all covariates, trustworthiness has the largest and most significant effect at 0.249 standard deviation units. Taken together, the results suggest that specific investigator characteristics — namely trustworthiness, followed by attractiveness and quality — explain at least part of the effect of racial concordance.

For completeness, we conduct the same exercise for sex concordance (Appendix Figure B4) despite not finding significant sex concordance effects in Model 0. While additions of investigator trustworthiness (Model 1), attractiveness (Model 2), and education (Model 4) flip the sign of the treatment coefficient, it remains small and insignificant across all specifications. We cannot reject the null that the treatment effect is the same across all models. Across all models, perceived trustworthiness is a positive, statistically significant, and robust predictor of trial participation intent (Appendix Tables A8 and A9).

We use the method proposed in Gelbach (2016) to further explore the mechanisms underlying the effect of racial concordance on interest in clinical trial participation. The method builds on the omitted variable bias formula and decomposes the change in the main coefficient between the base model (Model 0) and a version of the model including the variables that are candidates for

[3] See Appendix Table A7 for a correlation matrix of investigator features.

Table 3
Decomposition of the Effect of NIH Investigator Racial Concordance on Clinical Trial Participation Interest.

	Model 0 (1)	Model 6 (2)	Explained (3)
Racial Concordance	0.350	0.213	0.137
Trustworthiness	Not Included	Included	0.077
Attractiveness	Not Included	Included	0.067
Quality	Not Included	Included	0.023
Age	Not Included	Included	−0.026
Education	Not Included	Included	−0.004

Notes: OLS estimates of the willingness to participate in a future clinical trial led by race concordant NIH investigators. Outcomes are standardized to have a mean of 0 and standard deviation of 1. Column (1) includes an indicator for race concordant treatment. Column (2) includes all perceived investigator characteristics as defined in text. Column (3) shows the difference between the effect of racial concordance in Column (1) and Column (2) in the first row. The remaining rows show a decomposition of this difference into changes driven by the respective covariates added in Column (2). The decomposition follows the approach proposed in Gelbach (2016). All specifications include respondent characteristics (age, squared age, sex, education, employment, and South/non-South region) and ML-rated image characteristics (predicted investigator emotion, age, image quality, image size).

explaining the effect (Model 6). The coefficient on racial concordance in Model 0 is 0.350, which is reduced to 0.214 in Model 6. Therefore, the difference that is to be explained is 0.136. Table 3 shows that trustworthiness explains 0.077 units of the change, attractiveness explains 0.067 units, and quality explains 0.023 units. Including the age and education variables actually increased the portion of the racial concordance effect that is not explained by the included characteristics. This confirms that trustworthiness and attractiveness seem to be main drivers of the racial concordance effect with trustworthiness explaining 22 percent and attractiveness 19.1 percent of its effect (calculated by dividing the explained parts by the main coefficients in Model 0).

4.2. Robustness tests

We conduct several robustness checks to confirm the results. First, we conduct a further check for the presence of experimenter demand effects. To do this, we re-estimate the main effects, excluding the individuals who mentioned race- or sex-related terms in the open-text question about the study's purpose. The results, reported in Appendix Table A10, show similar effects. A second concern is that results are driven by outlier images — this is a legitimate concern given the few images of Black investigators we have based on the limited availability on the NIH website. We address this concern in two ways. First, we show the estimates from regressing interest in clinical trial participation on investigator fixed effects and our standard set of (respondent and image) controls in Appendix Figure B5. We cannot reject the null that the coefficients for all racially concordant investigators are the same (F-statistic = 1.27, p-value = 0.235) and that the coefficients for all discordant investigators are the same (F-statistic = 1.39, p-value = 0.143). Second, we rerun our main specification excluding each investigator in turn and report the distribution of estimates in Appendix Figure B6. These estimates range between 0.310 and 0.505 and most of them are close to our baseline estimate. We conclude that our estimated effect of racial concordance cannot be explained by a single outlier investigator. An additional concern is that our results are sensitive to the coding of our outcome variable. To address this concern, we demonstrate that our results are robust to alternative definitions of the outcome variable in Appendix Table A11. Lastly, we check whether our estimated effect is robust to controlling for other variables that may explain trial participation, including sex concordance with the investigator, having a medical condition, risk aversion, trust in others, and altruism. Appendix Table A12 shows the results when we include these controls. We fail to reject the null that the racial concordance estimates are the same as our baseline estimates when controlling for sex concordance (p-value = 0.947) and sex concordance as well as other respondent characteristics (p-value = 0.794).

4.3. Limitations and applicability of our findings

Randomized evaluations have many strengths but also well-known weaknesses — such as weak applicability to real-world situations. In our study, we presented participants with a picture of an NIH investigator and elicited their interest in participating in a clinical trial led by the investigator. In practice, featuring investigator images during trial recruitment is uncommon. However, it is increasingly common to highlight the contribution of Black investigators when aiming to engender trust, boost Black enrollment in trials and/or take-up of novel medical products (*e.g.*, Dr. Kizzmekia Corbett and the COVID-19 vaccine) (CBS News, 2021). When considering extrapolation to other settings, it is important that all of the images shown in our study were of actual NIH investigators and that subjects were informed of this fact. Results may differ if images were of investigators in a less-reputable institution or if the institution was not specified; previous research has shown lack of information increases opportunities for implicit bias to play a role in decision-making (Doleac and Hansen, 2020; Laouénan and Rathelot, 2022).

An additional potential concern with our findings is that we used a stated instead of a revealed preference measure. Enrolling subjects in actual medical experiments was not feasible as it would have been logistically challenging given the limited number of Black investigators and the specific enrollment criteria for each trial. Yet, given the importance of clinical trial participation to biomedical research productivity *and* the fact that it requires informed consent, stated willingness to participate is often used as an

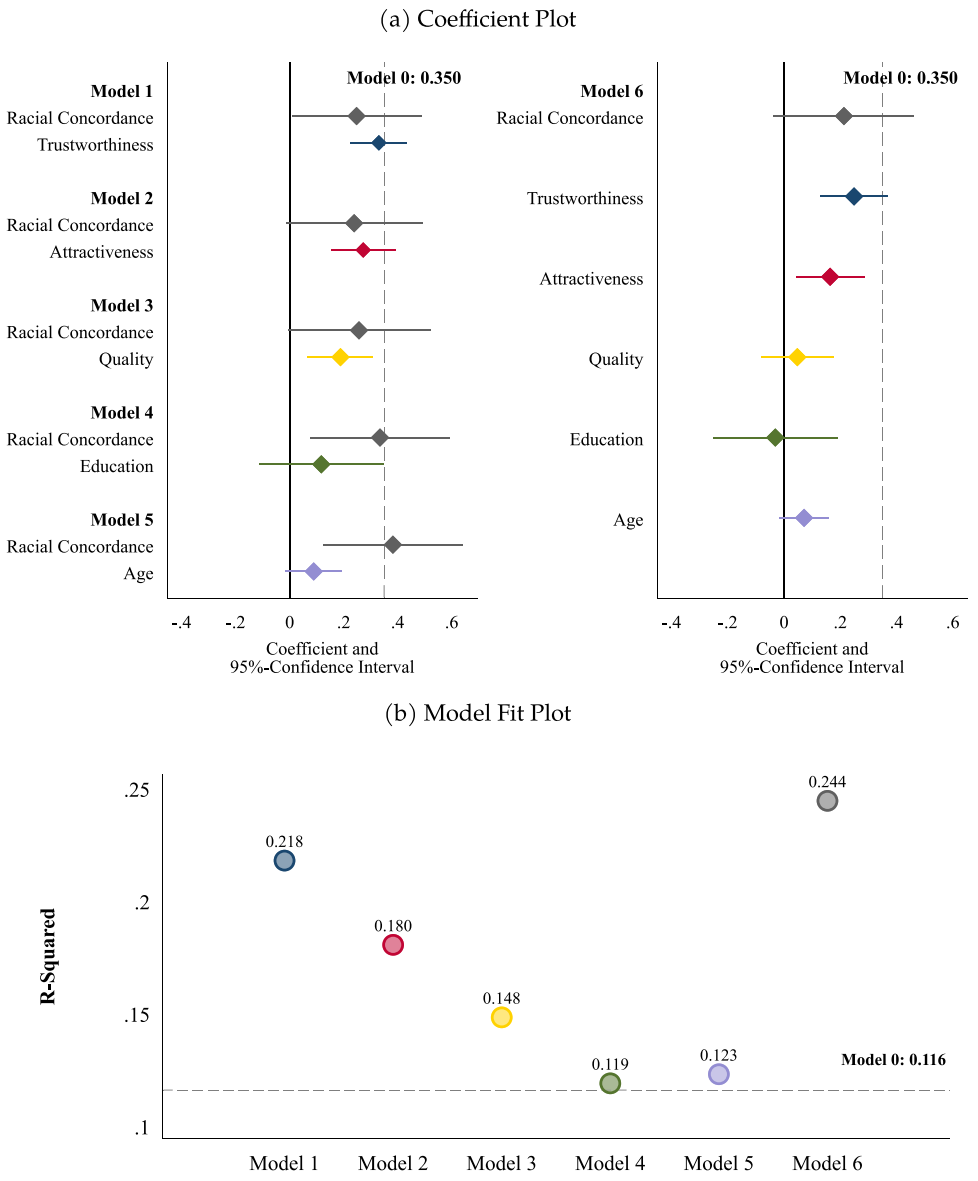

Fig. 2. The effect of NIH investigator racial concordance and perceived characteristics on clinical trial participation interest.
Notes: Figure plots OLS estimates from the six separate models described in the text. The outcome across all specifications is the willingness to participate in a clinical study led by the investigator, which is standardized to have a mean of 0 and standard deviation of 1. Panel (A) plots estimates for specifications that include the race concordant treatment and a *single* perceived investigator characteristic. Panel (B) plots estimates for a model with race concordant treatment and *all* perceived investigator characteristics. All specifications include respondent characteristics (age, squared age, sex, education, employment, and South/non-South region) and ML-rated image characteristics (predicted investigator emotion, age, image quality, image size). 95% CIs are plotted and based on robust standard errors. Panel (C) plots the R-squared value for each respective model. Dashed lines represent the estimated racial concordance effect (Panels A and B) and R-squared value (Panel C) for our main specification (Model 0, Table 1, Panel A, Column (4)).

outcome (*e.g.*, Moorcraft et al. (2016) and Pariera et al. (2017)). We searched the term "willingness to participate in clinical trials" in PubMed — and found that it has been the subject of more than 4500 publications in PubMed since 1990, many of which use stated "willingness to participate" as an outcome.[4] The number of publications with this key phrase has steadily trended upward over time (Appendix Figure B7).

[4] PubMed search results for "willingness to participate in clinical trials" can be accessed here. It is also the case that information and logistical hurdles pose key challenges to participation though are not our focus herein (Unger et al., 2020).

Insights from behavioral economics research support the use of stated preferences as proxies for revealed preference measures. Levitt and List (2007) show there can be a strong correlation between lab and real-world decisions. To assess these correlations in our setting, we leverage data collected by a nonprofit, Research!America, which annually surveys the U.S. public on their views regarding science and technology. We take advantage of a specific wave (2021) when the survey asked both about willingness to participate in a clinical trial as well as whether the respondent (or their family member) had enrolled in a trial in the past. The results, shown in Appendix Figure B8, demonstrate that 43 percent of respondents who stated that they were "very likely" to participate in a future clinical trial actually had participated in the past, compared to only 4 percent of respondents who stated they "would not participate".

To further validate our use of stated willingness to participate, we also analyzed alternative outcomes that are plausibly associated with future clinical trial participation. First, we asked respondents, "Would you like to learn more about research studies led by this National Institutes of Health Scientist?" directly after our intervention and primary outcome elicitation. In addition, at the end of the survey, we allowed respondents to act by clicking a link that directed to them more information on clinical trial enrollment. We also added a question on publicly financing research adopted from Research!America: "Would you be willing to pay $1 per week more in taxes if you were certain that all of the money would be spent on additional medical research?" We find that the effect of racial concordance on these three alternative measures is generally positive but small, and insignificant (Appendix Table A13).[5] This is likely because the outcomes were dichotomous, thus had low power, and were only indirectly related to the investigator treatment.

5. Discussion

In a randomized survey experiment using photographs of NIH investigators, Black respondents' interest in clinical trial participation increased when presented with a photograph of a Black investigator. We find that perceived investigator trustworthiness followed by attractiveness are the most important factors predicting interest in clinical trial participation using regression analysis and the Gelbach (2016) decomposition. We do not observe any effect of sex concordance on interest in participation. These findings suggest that racial concordance among principal investigators and potential research participants can improve enrollment of underrepresented racial minorities in medical research by projecting trustworthiness.

Our results are consistent with other literature emphasizing trustworthiness as a significant determinant of participation in the healthcare system (Jacobs et al., 2006; Warren et al., 2020). The historical and contemporaneous mistreatment and neglect of Black Americans, including but not limited to the infamous Tuskegee syphilis experiment, have negatively affected Black Americans' trust in the medical profession (Alsan and Wanamaker, 2018). A 2021 Pew Research Center survey found that the majority of Black adults think that research misconduct is just as likely to occur today as it was in the past (Funk, 2022). Further, in a July 2017 survey comparing attitudes about clinical trials among White, Black, Asian, and Hispanic U.S. adults, 50 percent of Black respondents cited lack of trust as a barrier to participating in clinical trials and 50 percent believed that subjects are sometimes included in clinical trials without being told (Research!America, 2017). Increasing trust among Black communities requires the medical profession to become more trustworthy (Warren et al., 2020). Part of building trust is increasing the opportunity for members of underrepresented groups to be in positions of authority, including as principal investigators and physicians.

While better representation of Black and other historically marginalized groups may be an effective tool for encouraging participation in clinical studies, it is not likely to be a comprehensive solution given the structural barriers to accessing medical care and obtaining information on trial opportunities as well as the time and financial burden imposed on participants (National Academies of Sciences, Engineering, and Medicine, 2022). Further research is needed to determine the importance of investigator characteristics relative to these other factors and to assess potential complementarities.

CRediT authorship contribution statement

Marcella Alsan: Writing – review & editing, Writing – original draft. **Romaine A. Campbell:** Writing – original draft. **Lukas Leister:** Formal analysis, Data curation. **Ayotomiwa Ojo:** Writing – review & editing, Writing – original draft.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jhealeco.2025.102968.

[5] The three outcomes are also significantly positively correlated with our primary outcome, stated willingness to participate, validating data quality and consistency of respondents' preferences.

References

Alegria, Mikhaila, Sud, Shawn, Steinberg, Benjamin E., Gai, Nan, Siddiqui, Asad, 2021. Reporting of participant race, sex, and socioeconomic status in randomized clinical trials in general medical journals, 2015 vs 2019. JAMA Netw. Open 4 (5), e2111516.

Alsan, Marcella, Durvasula, Maya, Gupta, Harsh, Schwartzstein, Joshua, Williams, Heidi L., 2024. Representation and extrapolation: Evidence from clinical trials. Q. J. Econ. 139 (1), 575–635, .

Alsan, Marcella, Eichmeyer, Sarah, 2024. Experimental evidence on the effectiveness of non-experts for improving vaccine demand. Am. Econ. J. Econ. Policy 16 (1), 394–414.

Alsan, Marcella, Garrick, Owen, Graziani, Grant, 2019. Does diversity matter for health? Experimental evidence from Oakland. Amer. Econ. Rev. 109 (12), 4071–4111.

Alsan, Marcella, Stanford, Fatima Cody, Banerjee, Abhijit, Breza, Emily, Chandrasekhar, Arun G., Eichmeyer, Sarah, Goldsmith-Pinkham, Paul, Ogbu-Nwobodo, Lucy, Olken, Benjamin A., Torres, Carlos, et al., 2021. Comparison of knowledge and information-seeking behavior after general COVID-19 public health messages and messages tailored for black and latinx communities: a randomized controlled trial. Ann. Intern. Med. 174 (4), 484–492.

Alsan, Marcella, Wanamaker, Marianne, 2018. Tuskegee and the health of black men. Q. J. Econ. 133 (1), 407–455.

Anderson, Steven R., Gianola, Morgan, Perry, Jenna M., Losin, Elizabeth A. Reynolds, 2020. Clinician–Patient racial/ethnic concordance influences racial/ethnic minority pain: Evidence from simulated clinical interactions. Pain Med. 21 (11), 3109–3125.

Artiga, Samantha, Kates, Jennifer, Michaud, Josh, Hill, L., 2022. Racial Diversity Within COVID-19 Vaccine Clinical Trials: Key Questions and Answers. Kaiser Family Foundation, https://www.kff.org/racial-equity-and-health-policy/issue-brief/racial-diversity-within-covid-19-vaccine-clinical-trials-key-questions-and-answers/. (Accessed 7 October 2022).

Awidi, Muhammad, Al Hadidi, Samer, 2021. Participation of black Americans in cancer clinical trials: Current challenges and proposed solutions. JCO Oncol. Pr. 17 (5), 265–271.

Banerjee, Abhijit, Alsan, Marcella, Breza, Emily, Chandrasekhar, Arun G., Chowdhury, Abhijit, Duflo, Esther, Goldsmith-Pinkham, Paul, Olken, Benjamin A., 2020. Messages on COVID-19 prevention in India increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities. NBER Working Paper No. 27496, National Bureau of Economic Research.

Basu, Anirban, Gujral, Kritee, 2020. Evidence generation, decision making, and consequent growth in health disparities. Proc. Natl. Acad. Sci. 117 (25), 14042–14051.

Bertrand, Marianne, Karlan, Dean, Mullainathan, Sendhil, Shafir, Eldar, Zinman, Jonathan, 2010. What's advertising content worth? Evidence from a consumer credit marketing field experiment. Q. J. Econ. 125 (1), 263–306.

Brandt, Allan M., 1987. No Magic Bullet: A Social History of Venereal Disease in the United States Since 1880. Oxford University Press, USA.

CBS News, 2021. Meet the black female scientist at the forefront of COVID-19 vaccine development. https://www.cbsnews.com/news/covid-19-vaccine-development-kizzmekia-corbett/. (Accessed 9 June 2023).

Chandler, Jesse, Rosenzweig, Cheskie, Moss, Aaron J., Robinson, Jonathan, Litman, Leib, 2019. Online panels in social science research: Expanding sampling methods beyond mechanical turk. Behav. Res. Methods 51, 2022–2038.

Chu, James, Pink, Sophia L., Willer, Robb, 2021. Religious identity cues increase vaccination intentions and trust in medical experts among American christians. Proc. Natl. Acad. Sci. 118 (49), e2106481118.

Cooper, Lisa A., Roter, Debra L., Johnson, Rachel L., Ford, Daniel E., Steinwachs, Donald M., Powe, Neil R., 2003. Patient-centered communication, ratings of care, and concordance of patient and physician race. Ann. Intern. Med. 139 (11), 907–915.

Deming, David J., Yuchtman, Noam, Abulafi, Amira, Goldin, Claudia, Katz, Lawrence F., 2016. The value of postsecondary credentials in the labor market: An experimental study. Amer. Econ. Rev. 106 (3), 778–806.

Doleac, Jennifer L., Hansen, Benjamin, 2020. The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden. J. Labor Econ. 38 (2), 321–374.

FDA, 2014. FDA action plan to enhance the collection and availability of demographic subgroup data. https://www.fda.gov/media/89307/download. (Accessed 7 October 2022).

FDA, 2022. Diversity plans to improve enrollment of participants from underrepresented racial and ethnic populations in clinical trials; draft guidance for industry. https://www.regulations.gov/document/FDA-2021-D-0789-0002. (Accessed 7 October 2022).

Frakes, Michael D., Gruber, Jonathan, 2022. Racial concordance and the quality of medical care: Evidence from the military. NBER Working Paper No. 30767, National Bureau of Economic Research.

Funk, Cary, 2022. Black Americans' trust in medical scientists and views about the potential for researcher misconduct. https://www.pewresearch.org/science/2022/04/07/black-americans-trust-in-medical-scientists-and-views-about-the-potential-for-researcher-misconduct/. (Accessed 7 October 2022).

Gelbach, Jonah B., 2016. When do covariates matter? And which ones, and how much? J. Labor Econ. 34 (2), 509–543, URL https://www.jstor.org/stable/26553211.

Ginther, Donna K., Schaffer, Walter T., Schnell, Joshua, Masimore, Beth, Liu, Faye, Haak, Laurel L., Kington, Raynard, 2011. Race, ethnicity, and NIH research awards. Science 333 (6045), 1015–1019.

Green, Angela K., Trivedi, Niti, Hsu, Jennifer J., Yu, Nancy L., Bach, Peter B., Chimonas, Susan, 2022. Despite the fda's five-year plan, black patients remain inadequately represented in clinical trials for drugs: Study examines FDA's five-year action plan aimed at improving diversity in and transparency of pivotal clinical trials for newly-approved drugs. Health Aff. 41 (3), 368–374.

Greenwood, Brad N., Carnahan, Seth, Huang, Laura, 2018. Patient–physician gender concordance and increased mortality among female heart attack patients. Proc. Natl. Acad. Sci. 115 (34), 8569–8574.

Greenwood, Brad N., Hardeman, Rachel R., Huang, Laura, Sojourner, Aaron, 2020. Physician–patient racial concordance and disparities in birthing mortality for newborns. Proc. Natl. Acad. Sci. 117 (35), 21194–21200.

Hill, Andrew, Jones, Daniel, Woodworth, Lindsey, 2020. Physician-patient race-match reduces patient mortality. Available at SSRN 3211276.

Jacobs, Elizabeth A., Rolle, Italia, Ferrans, Carol Estwing, Whitaker, Eric E., Warnecke, Richard B., 2006. Understanding African Americans' views of the trustworthiness of physicians. J. Gen. Intern. Med. 21, 642–647.

Laouénan, Morgane, Rathelot, Roland, 2022. Can information reduce ethnic discrimination? Evidence from AirBnB. Am. Econ. J. Appl. Econ. 14 (1), 107–132.

Lau, Emily S., Hayes, Sharonne N., Volgman, Annabelle Santos, Lindley, Kathryn, Pepine, Carl J., Wood, Malissa J., American College of Cardiology Cardiovascular Disease in Women Section, 2021. Does patient-physician gender concordance influence patient perceptions or outcomes? J. Am. Coll. Cardiol. 77 (8), 1135–1138.

Laurencin, Cato T., Jones, Camara P., Holden, Lynne M., National Academies of Sciences, Engineering, and Medicine, et al., 2023. Anti-black racism and the mental health of black men and black women in STEM. In: The State of Anti-Black Racism in the United States: Reflections and Solutions from the Roundtable on Black Men and Black Women in Science, Engineering, and Medicine: Proceedings of a Workshop. National Academies Press, Washington, DC.

Levitt, Steven D., List, John A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? J. Econ. Perspect. 21 (2), 153–174, .

Loeb, Stacy, Ravenell, Joseph E., Gomez, Scarlett Lin, Borno, Hala T., Siu, Katherine, Nolasco, Tatiana Sanchez, Byrne, Nataliya, Wilson, Godfrey, Griffith, Derek M, Crocker, Rob, et al., 2023. The effect of racial concordance on patient trust in online videos about prostate cancer: a randomized clinical trial. JAMA Netw. Open 6 (7), e2324395.

Ludwig, Jens, Mullainathan, Sendhil, 2023. Machine learning as a tool for hypothesis generation. NBER Working Paper No. 31017, National Bureau of Economic Research.

Manski, Charles F., Mullahy, John, Venkataramani, Atheendar S., 2023. Using measures of race to make clinical predictions: Decision making, patient health, and fairness. Proc. Natl. Acad. Sci. 120 (35), e2303370120.

Moorcraft, Sing Yu, Marriott, Cheryl, Peckitt, Clare, Cunningham, David, Chau, Ian, Starling, Naureen, Watkins, David, Rao, Sheela, 2016. Patients' willingness to participate in clinical trials and their views on aspects of cancer research: Results of a prospective patient survey. Trials 17 (1), 1–12.

National Academies of Sciences, Engineering, and Medicine, 2022. Recommendations for improving representation in clinical trials and clinical research. In: Bibbins-Domingo, Kristen, Helman, Alex (Eds.), Improving Representation in Clinical Trials and Research: Building Equity for Women and Underrepresented Groups. National Academies Press, Washington, DC, p. 133.

National Academies of Sciences, Engineering, and Medicine, 2023. The gatekeepers of STEMM: How individual bias and inequality persist and how STEM professionals can help. In: Barabino, Gilda A., Fiske, Susan T., Scherer, Layne A., Vargas, Emily A. (Eds.), Advancing Antiracism, Diversity, Equity, and Inclusion in STEMM Organizations: Beyond Broadening Participation. National Academies Press, Washington, DC, p. 165.

Nazha, Bassel, Mishra, Manoj, Pentz, Rebecca, Owonikoko, Taofeek K., 2019. Enrollment of racial minorities in clinical trials: Old problem assumes new urgency in the age of immunotherapy. Am. Soc. Clin. Oncol. Educ. Book 39, 3–10.

NIH, 2022a. FAQs about research supplements to promote diversity in health-related research (diversity supplements). https://nida.nih.gov/about-nida/organization/offices/office-nida-director-od. (Accessed 7 February 2022).

NIH, 2022b. Number of NIH principal investigators funded by grant mechanism and race. https://report.nih.gov/nihdatabook/report/306. (Accessed 7 August 2022).

NIH, 2022c. Principal investigators. myehosthttps://irp.nih.gov/our-research/principal-investigators/name. (Accessed 7 October 2022).

NIH, 2023. Racial Disparities in NIH Funding. National Institutes of Health, https://diversity.nih.gov. (Accessed 7 February 2022).

NIMHD, 2022. National institute of minority health and health disparities: History. https://www.nimhd.nih.gov/about/overview/history/. (Accessed 7 October 2022).

Ocampo, Ricardo, 2022. Image-quality 1.2.7. https://pypi.org/project/image-quality/. (Accessed 30 June 2021).

Pariera, Katrina L., Murphy, Sheila T., Meng, Jingbo, McLaughlin, Margaret L., 2017. Exploring willingness to participate in clinical trials by ethnicity. J. Racial Ethn. Heal. Disparities 4, 763–769.

Reardon, Sara, 2023. Alzheimer's drug trials plagued by lack of racial diversity. Nature 620 (7973), 256–257.

Research!America, 2017. National public opinion survey, clinical trials: Survey data ofl minority populations. https://www.researchamerica.org/wp-content/uploads/2022/07/July2017ClinTrialMinorityOversamplesPressReleaseSlidesFINAL_0-1.pdf. (Accessed 7 October 2022).

Roberts, Dorothy, 1997. Killing the Black Body: Race, Reproduction, and The Meaning of Liberty. Vintage Books, New York.

Scharff, Darcell P., Mathews, Katherine J., Jackson, Pamela, Hoffsuemmer, Jonathan, Martin, Emeobong, Edwards, Dorothy, 2010. More than tuskegee: Understanding mistrust about research participation. J. Health Care Poor Underserved 21 (3), 879.

Serengil, Sefik Ilkin, Ozpinar, Alper, 2021. Hyperextended lightface: A facial attribute analysis framework. In: 2021 International Conference on Engineering and Emerging Technologies. ICEET, pp. 1–4.

Street, Richard L., O'Malley, Kimberly J., Cooper, Lisa A., Haidet, Paul, 2008. Understanding concordance in patient-physician relationships: personal and ethnic dimensions of shared identity. Ann. Fam. Med. 6 (3), 198–205.

Takeshita, Junko, Wang, Shiyu, Loren, Alison W., Mitra, Nandita, Shults, Justine, Shin, Daniel B., Sawinski, Deirdre L., 2020. Association of racial/ethnic and gender concordance between patients and physicians with patient experience ratings. JAMA Netw. Open 3 (11), e2024583.

Unger, Joseph M., Hershman, Dawn L., Till, Cathee, Minasian, Lori M, Osarogiagbon, Raymond U., Fleury, Mark E., Vaidya, Riha, 2020. "When Offered to Participate": A Systematic Review and Meta-Analysis of Patient Agreement to Participate in Cancer Clinical Trials. JNCI: J. Natl. Cancer Inst. 113 (3), 244–257, .

Warren, Rueben C., Forrow, Lachlan, Hodge, Sr., David Augustin, Truog, Robert D., 2020. Trustworthiness before trust — Covid-19 vaccine trials and the black community. N. Engl. J. Med. 383 (22), e121.

Washington, Harriet A., 2006. Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present. Doubleday, New York.