DR. RACHEL EMILY SOLNICK (Orcid ID : 0000-0002-2565-5690)

Article type : Original Contribution

Emergency Physicians and Personal Narratives Improve the Perceived Effectiveness of COVID-19 Public Health Recommendations on Social Media: A Randomized Experiment

Rachel E. Solnick MD, MSc,[1,2] Grace Chao MD, MSc,[3,1,4] Ryan Ross MS,[5] Gordon T. Kraft-Todd PhD,[6] Keith E. Kocher MD, MPH [2,7,8]

[1]*U-M National Clinical Scholars Program*, Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI

[2]Department of Emergency Medicine, School of Medicine, University of Michigan; Ann Arbor, MI

[3] Veterans Affairs HSR&D Center for the Study of Healthcare Innovation, Implementation, & Policy, Ann Arbor, Michigan

[4] Department of Surgery, Yale School of Medicine, New Haven, Connecticut

[5]University of Michigan School of Public Health, Ann Arbor, Michigan

[6]Department of Psychology, Boston College, Chestnut Hill, Massachusetts

[7]Department of Learning Health Sciences, School of Medicine, University of Michigan; Ann Arbor, Michigan

[8]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, Michigan

Corresponding author and request for reprints:

Rachel E. Solnick, MD, MSc

Clinical Lecturer, Department of Emergency Medicine

National Clinician Scholars Program, Institute for Healthcare Policy and Innovation
University of Michigan
2800 Plymouth Rd, Bldg 14, Room G100
Ann Arbor, MI 48109
(832) 443-0512
solnick@umich.edu

Trial Registration: https://osf.io/kabvr (embargoed prior to publication)

Running Title: Emergency Physicians and Personal Narratives Improve the Perceived Effectiveness of COVID-19 Recommendations on Twitter

Keywords: COVID-19; social media; public health messaging; Twitter; randomized experiment; survey experiment

Word Count: 4978
Prior Presentations: None

Financial Support: The University of Michigan *National Clinical Scholars Program* supported this work. Dr Solnick was supported by the Institute for Healthcare Policy and Innovation at U-M National Clinician Scholars Program. Dr. Chao was supported by the VA Office of Academic Affiliations through the VA/National Clinician Scholars Program and the University of Michigan Medicine at UM. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

Author Contributions: Dr Solnick had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Solnick, Chao, Kraft-Todd, Ross, Kocher.
Acquisition, analysis, or interpretation of data: Solnick, Chao, Kraft-Todd, Ross, Kocher.
Drafting of the manuscript: Solnick, Chao, Kraft-Todd, Ross, Kocher.
Critical revision of the manuscript for important intellectual content: Solnick, Chao, Kraft-Todd, Ross, Kocher.
Statistical analysis: Solnick, Kraft-Todd, Ross
Obtained funding: Solnick.

Administrative, technical, or material support: Solnick, Kocher.

Supervision: Kocher.

Conflict of Interest Disclosures: No authors reported disclosures.

Role of the Funder/Sponsor: The University of Michigan *National Clinical Scholars Program* provided funding for the study but had no role in the conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; the decision to submit the manuscript for publication; or the decision as to which journal the manuscript was submitted. No other funding sources had a role in the study.

Abstract

Background: Containment of the coronavirus disease 2019 (COVID-19) pandemic requires the public to change behavior under social distancing mandates. Social media are important information dissemination platforms that can augment traditional channels communicating public health recommendations. The objective of the study is to assess the effectiveness of COVID-19 public health messaging on Twitter when delivered by emergency physicians and containing personal narratives.

Methods: On April 30, 2020, we randomly assigned 2007 U.S. adults to an online survey using a 2x2 factorial design. Participants rated 1 of 4 simulated Twitter posts varied by messenger type (emergency physician vs federal official) and content (personal narrative vs impersonal guidance). Main outcomes were: perceived message effectiveness (35-point scale); perceived attitude effectiveness (15-point scale); likelihood to share Tweets (7-point scale); and writing a letter to their governor to continue COVID-19 restrictions (write letter or none).

Results: The physician/personal message had the strongest effect and significantly improved all main messaging outcomes except for letter-writing. Unadjusted mean differences between physician/personal and federal/impersonal were: perceived messaging effectiveness (3.2 [95%CI, 2.4-4.0]); perceived attitude effectiveness (1.3 [95%CI, 0.8-1.7]); likelihood to share

(0.4 [95%CI, 0.15-0.7]). For letter-writing, physician/ personal made no significant impact compared to federal/ impersonal (odds ratio 1.14 [95%CI, 0.89-1.46]).

Conclusions: Emergency physicians sharing personal narratives on Twitter are perceived to be more effective at communicating COVID-19 health recommendations compared to federal officials sharing impersonal guidance.

INTRODUCTION

The coronavirus disease 2019 (COVID-19) crisis has exposed the critical need for clearly and consistently communicating public health guidelines anchored in the best available evidence. Yet, many voices are competing with public health officials, particularly given that social media outlets frequently supplant traditional news sources.[1] Amid this backdrop, the U.S. has had higher COVID-19-associated deaths and excess all-cause mortality compared to most peer countries.[2] Despite the alarming rate of viral transmission, the public has not had full compliance with pandemic guidelines.[3,4] Policymakers and public health officials therefore must be strategic in communicating pandemic-related messages to the public.

Emergency physicians can play a key role in disseminating and amplifying public health recommendations especially during a crisis.[5,6] Emergency departments experienced the severity of the initial COVID-19 viral surge and were challenged by a rapid response to the influx of ED patients.[7–9] Serving at the front lines of the epidemic, emergency physicians have played a prominent role as a trusted source in communicating COVID-19 updates and urging the public to stay home.[6,10,11] The effectiveness of public messaging can be influenced by the credibility of the messenger [12,13] and the content of the message.[14] However, there is little experimental data measuring the effectiveness of public health communication through personal narrative or by physicians, which has been commonly seen in social media posts during the COVID-19 pandemic.

Therefore, the goal of this study was to evaluate the effectiveness of a physician versus federal official and personal versus impersonal content in delivering COVID-19 public health recommendations on Twitter, a popular social media platform. We tested the following hypotheses: 1) Emergency physicians deliver a more effective message than federal officials; 2) Personal appeals are more effective than impersonal ones; and 3) The interaction of a physician messenger with a personal message is synergistic.

METHODS

Study Design and Setting

We conducted a preregistered randomized experiment using simulated Twitter accounts and posts that randomly manipulated messenger type and message content in a 2 × 2 between-subject factorial design. We launched the experiment on April 30, 2020, the day the White House-issued public restriction guidelines were set to expire, transferring decision-making responsibility on restrictions to state governments.

This trial was approved by the institutional review board at the University of Michigan. Written informed consent was obtained from participants before participation. This trial followed the Consolidated Standards of Reporting Trials (CONSORT)[15] guideline with suggested amendments for reporting nonpharmacologial treatments and factorial trials.[16]

Participants

We recruited U.S.-based adult participants from Lucid Theorem, a nationally representative crowdsourced online subject pool that is quota-sampled to match census demographics on age, gender, race/ethnicity, and region.[17] Participants were eligible if ≥18 years old. We included responses for analysis if ≥80% of study questions were complete. We assessed the impact of weighting the sample based on demographic characteristics of U.S. adults with Internet access as reported by the 2017 U.S. Census.[18] **(eTable 1 and eTable 2)** Participants in Lucid were compensated at a rate comparable to $1 per study. Median time to complete the study was 11 minutes.

Study Procedures

Participants accessed the online survey (Qualtrics, Provo UT) through their personal electronic devices and gave consent blinded to the study objectives. They first underwent a pre-treatment attention assessment with the correct answer embedded in the instruction stem.[19] We randomized participants to 1 of 4 treatment arms with simulated Twitter posts and they answered a series of questions to measure primary outcomes. This was followed by a second attention check to recall the messenger's occupation which was a means of assessing that participants read the post and had received the intervention. Lastly, participants were invited to take a stay-at-home pledge, write a letter to their governor, and to answer additional covariate questions.

Twitter Stimuli and Randomization

We created images of a Twitter account and message for experimental exposures. We used the same male actor for the emergency physician (dressed in scrubs and a surgical cap) and the non-physician federal official (business clothes). The background photo was a building selected to plausibly appear as either a federal building or hospital. We took other Twitter metrics (date joined, number of accounts followed and followers) from an exemplar emergency physician Twitter account which were the same across conditions.

For message content, we compared the effect of a personal versus impersonal message. The personal message was based on "the identifiable victim effect", that having more identifiable information about a victim increases caring.[20] In contrast, the language for the impersonal message was used directly from a mass federal communication mailed on postcards to 130 million U.S. households[21] as part of the "President's Coronavirus Guidelines for America" and from the White House "Opening up America Again" guidelines.[22,23]

The two messages had approximately the same number of words (personal:61, impersonal:55) and delivered a similar three-part message: (1) young people are at risk; (2) public activity restrictions should continue; and (3) continuing restrictions would reduce the risk of viral resurgence. (**Figure 1**).

Simple random assignment was accomplished via the randomizer tool in Qualtrics. Each participant was assigned to 1 of 4 possible treatment arms with equal probability: 498 to physician/personal (PP); 505 to physician/impersonal (PI); 505 to federal/personal (FP), and 499 to federal/impersonal (FI).

Primary Outcome Measures

To evaluate the effect of messages, we measured (1) perceived message effectiveness (PME), (2) perceived attitude effectiveness (PAE), and (3) behavioral outcomes: likelihood to share, write a letter to a governor. The PME scale was intended to measure the message's emotional impact, and was adapted from a scale used in the context of smoking cessation research.[24] Participants evaluated the messages as: memorable, grabbed my attention, powerful, meaningful, and convincing on a 7-point Likert scale "Strongly disagree" to "Strongly agree" (coded 1-7), summed to a 35-point rating. (Supplement section 5) We modified the original scale by removing subscale "informative," due to COVID-19 information saturation. The modified scale demonstrates high reliability (α=.93) and an eigenvalue of 3.96 accounting for 79.2% of the variance, similar to the original scale reliability (α=.94) and eigenvalue of 4.22 accounting for 70% of the variance.

The perceived attitude effectiveness (PAE) scale was intended to measure the message's effect on attitudes, and was adapted from a scale used in smoking cessation research.[25] Participants evaluated whether the message (1) "Made me concerned about the health effects of lifting restrictions on public activity"; (2) "Made lifting restrictions less appealing"; (3) "Discourages me from supporting opening America up right now" on a 5-point Likert scale, "Not effective at all" to "Extremely effective" (coded 1-5), summed to a 15-point rating. The modified scale demonstrates high reliability (α=.88) and one-factor dimension that accounted for 81.3% of the variance, similar to the original scale reliability (α=.93) and a general factor that accounted for 82.6% of the variance.

We measured likelihood to share the Tweet as an estimator of the messages' behavioral impact. This was measured on a 7-point Likert scale "Extremely unlikely" to "Extremely likely" (coded 1-7). Self-reported willingness to share social media posts has previously been correlated with increased sharing in reality.[26]

Lastly, we asked participants whether they were interested in writing a letter to their state governor (yes/no). Participants who agreed were provided a free-text response box to write to the governor (not a form letter) and were truthfully informed we would send this letter anonymously, which we did via state government online communication forms. Because of the cognitive effort involved, the letter-writing task is less susceptible to desirability bias.[27]

Secondary Outcome Measures

As an exploratory outcome, we asked participants to take a pledge (yes/no) to stay inside to fight COVID-19. Pledging has been a popular way in the COVID-19 pandemic for concerned groups to encourage social distancing.[28] Prior research indicates that pledging to engage in prosocial behavior (*e.g.*, voting, environmental protection) has a small but significant effect on increasing the desired outcome.[29]

Covariate Measures

We incorporated additional variables in a covariate-adjusted model and to explore heterogeneous treatment effects using demographic information provided by Lucid (age, education, race/ethnicity, sex, household income, political party, state), which we supplemented with survey questions on overall health, marital status, population density, number in household, employment status, and political ideology. We also collected variables related to health behaviors, policy positions, and messaging receptiveness: anxiety about coronavirus, trust in federal officials and physicians,[30] economy vs public health trade-off,[31] political engagement,[32] consumption of media bias via AllSides rankings,[33] empathy (using the empathic concern subscale of the Brief Interpersonal Reactivity Index[34]), and news exposure frequency. Finally, we incorporated data on the extent of COVID-19 cases and restrictions based on the participant's state of residence (Supplement section 3).

Statistical Analysis

Sample size was determined from a pilot survey with 601 Lucid participants conducted two weeks prior and not included in the final study. We estimated with 438 participants per treatment arm (N =1752), the minimum detectable effect at 80% power using a 2-sided hypothesis test ($\alpha = .05$) is approximately 0.10 standardized units for a bivariate outcome difference of letter-writing.

The statistical analysis plan was pre-registered prior to data collection through the Open Science Framework (Supplement Section 9). We compared demographic characteristics and outcomes across groups by analysis of variance and T-Test for continuous variables and $\chi 2$ test and Z-test of proportions for categorical variables. As recommended for the accurate reporting of factorial studies,[16] we present three major comparisons: (1) 4-level treatment effects; (2) each factor pooled (messenger and message content); and (3) interaction between factors. Assumptions for each statistical test were evaluated using standard diagnostic tests and no major violations were found.

We estimated treatment effects using ordinary least-squares linear regression and logistic regression on the 4-level treatment factor, with federal impersonal as the omitted reference category. Regression models were covariate-adjusted to maximize the precision of estimated treatment effects. Covariates were selected by items expected to be associated with social distancing, then manually backward selected for inclusion based on the strength of the association with the outcome and Akaike information criterion (AIC) of the model fit: race/ethnicity, marital status, political party, gender, COVID-19 anxiety, news frequency, and economy vs public health trade-off. All models were assessed for violations of basic assumptions and no major violations were found. Participants with missing value for a variable were included with a missing data indicator for that variable.

We also examined whether subgroups of participants were affected differently by treatments using generalized random forest, a machine learning algorithm that estimates treatment effect heterogeneity as a function of each participant's covariate profile by nonparametric statistical estimation based on random forests.[35] Understanding how demographics may contribute to different responses to messaging can help in creating tailored content for specific groups at higher- risk for COVID-19.[4] Identifying these groups would create opportunities for audience segmentation - varying messaging strategies to address different groups - as demonstrated in climate science communication literature.[36] We assessed the effect heterogeneity specifically for PME because as an emotion-based rapid cognition, we hypothesized it would be more likely to be influenced by demographic profiles.[37] R version 3.5.2 (R Foundation for Statistical Computing) was used for statistical analyses, and the grf package was used for Causal Forests.[38]

RESULTS

Of 2090 participants who entered the survey, 2007 consented, were randomized, and completed the survey with ≥80% data (**eFigure 1**). All participants that were randomized were included in the analysis. Participants' mean age was 45 years (SD 16.7 years), 51% (n=1034) were female, 10.6% (n=214) were Black, and 11.6% (n=234) were Hispanic. Baseline characteristics and covariates were well-balanced across the four treatment arms (**Table 1, eTable 3**).

Main Outcomes

For the 4-level treatment results, participants rated PME, PAE and likelihood to share significantly higher in the physician/personal (PP) condition compared with the federal/impersonal (FI) condition, with largest effect on PME (**Figure 2**). Unadjusted estimated effects of PP versus FI are presented here with outcome means (**eTable 4**); remaining comparisons are shown in **eTable 5**. For the PME 35-point scale outcome, the means were: PP 28.52 (SD 6.81) versus FI 25.32 (SD 6.95) (difference 3.2 [95%CI, 2.37 to 4.02] $p<0.001$). For the PAE 15-point scale, the means were: PP 11.02 (SD 3.66) versus FI 9.77 (3.54) (difference 1.26 [95%CI, 0.81 to 1.7] $p<0.001$). For the likelihood to share 7-point scale, the means were: PP 4.99 (2.09) vs FI 4.59 (2.13); (difference 0.4 [95%CI, 0.15 to 0.66] p=0.003). There was no significant difference across treatment arms of letter-writing to the governor to continue public activity restrictions (odds ratio for PP compared to FI was 1.14 [95%CI, 0.89-1.46]). The proportion letter-writing was 50.6% for PP vs 47.3% for FI (difference 3.3% [95%CI, -3.1%to 9.7%] p=0.33). There was similarly no significant effect on the pledge to stay home secondary outcome: mean PP 90.6% vs FI 90.0% p=0.99. As expected, adjusted means had similar effect estimates with more precise confidence intervals (**eTable 6**).

The average effects of the messenger and message are presented in **eTable 7**. The pooled treatment effect of both personal content and physician messenger had a statistically significant impact on both PME and PAE. Cohen's D, a standardized measure of effect size, is presented here to facilitate comparing across different scales-- 0.2 is considered a small effect and 0.5 a medium effect.[39] The average personal content had a stronger effect compared to physician messenger for PME (0.40 [95%CI, 0.28 to 0.52] p<0.001 versus 0.25 [95%CI, 0.13 to 0.37] p<0.001) and PAE (0.22 [95%CI, 0.10 to 0.35] p<0.001 versus 0.16 [95%CI, 0.04 to 0.29] p=0.009), respectively. Conversely, personal content did not significantly increase likelihood to share, while the physician messenger retained a positive effect (0.17 [95%CI, 0.05 to 0.30] p=0.006). We found a negative interaction for PME such that physicians had an incrementally increased score compared to federal officials when presenting for the impersonal context, but less so for the personal narrative (-1.18 [95% CI, -2.35 to -0.02]; P=0.045). No significant interactions were found for the other primary outcomes.

Sensitivity Analysis Attention Check Question

We presented participants with two attention checks. Most participants passed the post-outcome measured manipulation check, correctly selecting the occupation in the Twitter profile (81.1%, n=1628). Far fewer passed the pre-exposure check in which the correct answer was hidden within the instruction paragraph (52.1%, n=1046). The groups were similar in treatment effects but had slightly stronger effects in the groups with higher levels of attention checks. (eTable 8, **eFigure 2**)

Treatment Effect Heterogeneity

We did not find significant heterogeneity in causal forest-estimated treatment effects of the personal message on PME. Causal forest was trained on many key variables, and test set predictions and CIs were assessed (**Figure 3**). While some patterns visually emerged among the variables specifically selected for graphical illustration based on hypothesised effect heterogeneity- political ideology, health status, age, and race/ ethnicity- all individual confidence intervals overlapped, coinciding with the null global test.

DISCUSSION

To our knowledge, this is the first large-scale, nationally representative, pre-registered, randomized experiment to directly estimate the effect of a physician versus federal official messenger and message content of simulated social media posts on individual perceptions, attitudes, and behavior. We found that public health messages delivered by physicians and personal messages elicited stronger emotions, greater changes in attitudes and an increased willingness to disseminate the message than when federal officials delivered impersonal messages. We did not observe differences in a stay-at-home pledge (which was near ceiling), nor in willingness to write a letter to the governor to continue restrictions. These findings suggest that to emergency physicians sharing personal stories on social media may be more effective in increasing general adherence to public health guidelines than federal officials sharing impersonal messages. Complementary communication campaigns are still needed to augment these recommendations in order to change pandemic related individual behavior.

Our study adds important findings of source effects and messaging content on a non-traditional communication platform during this public health crisis. We demonstrate that trusted messengers can alter opinions on contentious public policy issues consistent with prior experiments finding a medical scientist and physician increased support for antimicrobial resistance policy[12] and comparative effectiveness research,[13] respectively. The framing of health messages also matters. Similar to identifiable victim effect findings, we found enhanced emotional and attitudinal impact when the message was to help a single, identifiable person (i.e. the COVID-19 victim who was a friend) compared to the concept of helping the many, unidentifiable others.[20,40] Moreover, findings of increased public health messaging effectiveness from personal narratives is also supported by organ donation literature, which has shown that when viewers are more emotionally involved in a television narrative they were more likely to become organ donors if the show encouraged donation.[41]

We also assessed heterogeneous treatment effects to determine if there were distinct subpopulations which were impacted by the intervention differently, a finding which would be helpful for tailoring messaging for different groups. Despite a rigorous investigation harnessing machine learning tools, we found no significant impact of any participant characteristic, on the extent or direction of the message's impact, specifically examining political ideology, health status, age, and race/ ethnicity. Although we did not observe a differential impact of the emergency physician or federal official on lower income or minority participants, underserved populations may have lower trust in physicians than those included in our study,[42] and may interact with messages differently from our participants. Future research should examine how to most effectively communicate with underserved minority populations hardest hit by the pandemic.

Our results add to a growing body of research investigating the impact of social media platforms for public health communication. The majority of Twitter users cite it as a news source,[1] presenting an opportunity for health professionals to capitalize on this channel as an adjunct for reaching a broader segment of the public. Physicians, scientists, and health providers have played an increasing role on Twitter, using it to share personal communications[43] and engage with the public on health issues.[44] Relevant to a pandemic, Twitter has been identified as a tool for efficient information dissemination during emergency events[5] and in public health crises to communicate recommendations.[45] Our findings support the increased use of Twitter by healthcare professionals as a platform to communicate directly to the public.

While government mandated public activity restrictions and social distancing recommendations play a key role in preventing the spread of COVID-19, these interventions will be ineffective if the public is not willing to adhere to them. Social media based public messaging may help to improve the public's perception of these measures and thus adherence to health guidelines. However, during the pandemic, several U.S. healthcare institutions urged physicians not to make public appeals.[46–49] Our findings bolster policies that protect social media use by scientists and health providers to share public health communications directly to the public.

Limitations

This study has several limitations. First, the experimental design used a simulated Twitter message in the context of an online survey. Federal officials may be restricted on what they can communicate on social media using their official titles, but pilot data for this experiment showed most participants found the Twitter stimuli believable. It is possible that participants would react differently if they encountered these messages on the actual social media platform. However, participant likelihood to share a post has been shown to correlate highly with action in real life.[26] Furthermore, while the effects of user comments on social media were beyond the scope of this study, prior research has shown that user comments may have an additive effect on messaging impact,[50,51] though whether it will change reader behavior is unknown. Although we observed an increased willingness to share certain messages, we did not find differences in pledging to stay home nor writing a letter to the governor to maintain restrictions. It remains unclear if the impact of the messages would translate into real-life changes in compliance with social distancing measures. Second, though the participant pool matches U.S. demographics in most regards, our participants had higher educational attainment and lower proportion of Hispanic origin (approximately 15.4% of U.S. population with access to internet versus 11% in our study)[18] We weighted our sample to account for educational differences and still did not observe an appreciable impact on treatment effects (**eTable 3**). Further supporting generalizability, Lucid participants have exhibited behavioral experimental results similar to U.S. national probability samples.[17] Third, the high levels of reported anxiety created a likely ceiling effect for our outcomes. For PME, almost half of participants rated the message at 6 or above on a 7-point scale. Ceiling effects may have reduced sensitivity to determining differences by treatment, biasing results towards null. Lastly, we selected white males for the physician and federal official in the study, the most common demographic for both groups. It is possible that other race and genders of the Twitter messenger could have influenced subpopulations of this study differently than white males, however prior patient satisfaction simulation studies did not find differences by physician race or gender.[52]

CONCLUSION

Using a rigorous randomized experiment of a simulated Twitter message, we found that an emergency physician's Twitter message of a personal story and recommendation related to COVID-19 increased the attitudinal, emotional and willingness to share measures of impact compared to a federal official sharing impersonal guidance. These results underscore the advocacy role for physicians on social media in promoting public health recommendations. We did not find an impact on letter writing to their governor to support COVID-19 restrictions nor pledging to stay home. Future directions should explore the real-world impact of emergency physician public health tweets on measures of behavior change.

References

1. The Evolving Role of News on Twitter and Facebook [Internet]. Pew Research Center's Journalism Project. 2015 [cited 2020 May 4];Available from: https://www.journalism.org/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook/

2. Bilinski A, Emanuel EJ. COVID-19 and Excess All-Cause Mortality in the US and 18 Comparison Countries. JAMA [Internet] 2020;Available from: http://dx.doi.org/10.1001/jama.2020.20717

3. Park CL, Russell BS, Fendrich M, Finkelstein-Fox L, Hutchison M, Becker J. Americans' COVID-19 Stress, Coping, and Adherence to CDC Guidelines. J Gen Intern Med [Internet] 2020;35(8):2296–303. Available from: http://dx.doi.org/10.1007/s11606-020-05898-9

4. Block R Jr, Berg A, Lennon RP, Miller EL, Nunez-Smith M. African American Adherence to COVID-19 Public Health Recommendations. Health Lit Res Pract [Internet] 2020;4(3):e166–70. Available from: http://dx.doi.org/10.3928/24748307-20200707-01

5. Hughes AL, Palen L. Twitter adoption and use in mass convergence and emergency events. IJEM [Internet] 2009 [cited 2020 May 4];6(3/4):248. Available from: http://www.inderscience.com/link.php?id=31564

6. Gaeta C, Brennessel R. COVID-19: Emergency Medicine Physician Empowered to Shape Perspectives on This Public Health Crisis. Cureus [Internet] 2020;12(4):e7504. Available from: http://dx.doi.org/10.7759/cureus.7504

7. Rothfeld M, Sengupta S, Goldstein J, Rosenthal BM. 13 Deaths in a Day: An “Apocalyptic” Coronavirus Surge at an N.Y.C. Hospital [Internet]. The New York Times. 2020 [cited 2020 Sep 24];Available from: https://www.nytimes.com/2020/03/25/nyregion/nyc-coronavirus-hospitals.html

8. Feldman N, Lane R, Iavicoli L, et al. A snapshot of emergency department volumes in the “epicenter of the epicenter” of the COVID-19 pandemic. Am J Emerg Med [Internet] 2020;Available from: http://dx.doi.org/10.1016/j.ajem.2020.08.057

9. Whiteside T, Kane E, Aljohani B, Alsamman M, Pourmand A. Redesigning emergency department operations amidst a viral pandemic. Am J Emerg Med [Internet] 2020;38(7):1448–53. Available from: http://dx.doi.org/10.1016/j.ajem.2020.04.032

10. Chokshi DA, Katz MH. Emerging Lessons From COVID-19 Response in New York City. JAMA [Internet] 2020;323(20):1996–7. Available from: http://dx.doi.org/10.1001/jama.2020.7310

11. Editors E. 50 Experts to Trust in a Pandemic [Internet]. Medium Elemental. 2020 [cited 2020 Sep 24];Available from: https://elemental.medium.com/50-experts-to-trust-in-a-pandemic-fe58932950e7

12. Martin A, Gravelle TB, Baekkeskov E, Lewis J, Kashima Y. Enlisting the support of trusted sources to tackle policy problems: The case of antimicrobial resistance. PLoS One [Internet] 2019 [cited 2020 Mar 27];14(3):e0212993. Available from: https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0212993&type=printable

13. Gerber AS, Patashnik EM, Doherty D, Dowling CM. Doctor knows best: physician endorsements, public opinion, and the politics of comparative effectiveness research. J Health Polit Policy Law [Internet] 2014;39(1):171–208. Available from: http://dx.doi.org/10.1215/03616878-2395208

14. Davis KC, Duke J, Shafer P, Patel D, Rodes R, Beistle D. Perceived Effectiveness of Antismoking Ads and Association with Quit Attempts Among Smokers: Evidence from the Tips From Former Smokers Campaign. Health Commun [Internet] 2017;32(8):931–8. Available from: http://dx.doi.org/10.1080/10410236.2016.1196413

15. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med [Internet] 2008;148(4):295–309. Available from: http://dx.doi.org/10.7326/0003-4819-148-4-200802190-00008

16. McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. JAMA [Internet] 2003;289(19):2545–53. Available from: http://dx.doi.org/10.1001/jama.289.19.2545

17. Coppock A, McClellan OA. Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. Research & Politics [Internet] 2019;6(1):2053168018822174. Available from: https://doi.org/10.1177/2053168018822174

18. Broadband Adoption and Computer Use by year, state, demographic characteristics - Data.gov [Internet]. [cited 2020 Nov 19];Available from: https://catalog.data.gov/dataset/broadband-adoption-and-computer-use-by-year-state-demographic-characteristics

19. Berinsky AJ, Margolis MF, Sances MW, Warshaw C. Using screeners to measure respondent attention on self-administered surveys: Which items and how many? Political Science Research and Methods [Internet] [cited 2020 Apr 29];1–8. Available from: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/979A15EB14DBBF596D56032D0CBB4424/S2049847019000530a.pdf/div-class-title-using-screeners-to-measure-respondent-attention-on-self-administered-surveys-which-items-and-how-many-div.pdf

20. Small DA, Loewenstein G. Helping a Victim or Helping the Victim: Altruism and Identifiability. J Risk Uncertain [Internet] 2003;26(1):5–16. Available from: https://doi.org/10.1023/A:1022299422219

21. Stankiewicz K. US households are being mailed “President Trump”s Coronavirus Guidelines for America’ [Internet]. CNBC. 2020 [cited 2020 May 1];Available from: https://www.cnbc.com/2020/03/27/us-households-are-being-mailed-trumps-coronavirus-guidelines.html

22. Opening Up America Again | The White House [Internet]. The White House. [cited 2020 May 1];Available from: https://www.whitehouse.gov/openingamerica/

23. POTUS-Coronavirus-Guidelines_30-DAYS. Available from: https://www.whitehouse.gov/wp-content/uploads/2020/03/03.16.20_coronavirus-guidance_8.5x11_315PM.pdf

24. Davis KC, Nonnemaker J, Duke J, Farrelly MC. Perceived effectiveness of cessation advertisements: the importance of audience reactions and practical implications for media campaign planning. Health Commun [Internet] 2013;28(5):461–72. Available from: http://dx.doi.org/10.1080/10410236.2012.696535

25. Baig SA, Noar SM, Gottfredson NC, Boynton MH, Ribisl KM, Brewer NT. UNC Perceived Message Effectiveness: Validation of a Brief Scale. Ann Behav Med [Internet] 2019;53(8):732–42. Available from: http://dx.doi.org/10.1093/abm/kay080

26. Mosleh M, Pennycook G, Rand DG. Self-reported willingness to share political news articles in online surveys correlates with actual sharing on Twitter. PLoS One [Internet] 2020;15(2):e0228882. Available from: http://dx.doi.org/10.1371/journal.pone.0228882

27. Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Psychol [Internet] 1960;24:349–54. Available from: http://dx.doi.org/10.1037/h0047358

28. Sign the Petition [Internet]. Change.org. [cited 2020 May 18];Available from: https://www.change.org/p/new-yorkers-plea-from-ed-doctors-to-listen-to-covid-19-public-health-advise?recruiter=20768044&utm_source=share_petition&utm_medium=twitter&utm_campaign=psf_combo_share_abi&utm_term=psf_combo_share_initial&recruited_by_id=aeb355d0-f085-012f-725f-4040acce234c

29. Costa M, Schaffner BF, Prevost A. Walking the walk? Experiments on the effect of pledging to vote on youth turnout. PLoS One [Internet] 2018;13(5):e0197066. Available from: http://dx.doi.org/10.1371/journal.pone.0197066

30. Reinhart by RJ. Nurses Continue to Rate Highest in Honesty, Ethics [Internet]. Gallup.com. 2020 [cited 2020 May 1];Available from:

https://news.gallup.com/poll/274673/nurses-continue-rate-highest-honesty-ethics.aspx

31. Ballew, M., Bergquist, P., Goldberg, M., Gustafson, A., Kotcher, J., Marlon, J., Roess, A., Rosenthal, S., Maibach, E., & Leiserowitz, A. American Public Responses to COVID-19 [Internet]. Yale University and George Mason University.; 2020. Available from: https://climatecommunication.yale.edu/publications/american-public-responses-to-covid-19-april-2020/

32. Appendix A: Measures and scales [Internet]. Pew Research Center - U.S. Politics & Policy. 2018 [cited 2020 Apr 29];Available from: https://www.people-press.org/2018/04/26/appendix-a-measures-and-scales-2/

33. AllSides Media Bias Ratings [Internet]. AllSides. [cited 2020 Apr 29];Available from: https://www.allsides.com/media-bias/media-bias-ratings

34. Ingoglia S, Lo Coco A, Albiero P. Development of a Brief Form of the Interpersonal Reactivity Index (B-IRI). J Pers Assess [Internet] 2016;98(5):461–71. Available from: http://dx.doi.org/10.1080/00223891.2016.1149858

35. Athey S, Tibshirani J, Wager S. Generalized random forests. Ann Stat [Internet] 2019 [cited 2020 May 6];47(2):1148–78. Available from: https://projecteuclid.org/euclid.aos/1547197251

36. Maibach EW, Leiserowitz A, Roser-Renouf C, Mertz CK. Identifying like-minded audiences for global warming public engagement campaigns: an audience segmentation analysis and tool development. PLoS One [Internet] 2011;6(3):e17571. Available from: http://dx.doi.org/10.1371/journal.pone.0017571

37. Slovic P, Finucane ML, Peters E, MacGregor DG. The affect heuristic. Eur J Oper Res [Internet] 2007;177(3):1333–52. Available from: http://www.sciencedirect.com/science/article/pii/S0377221705003577

38. Tibshirani J, Athey S, Friedberg R. Package “Grf.” [Internet]. 2020. Available from: https://cran.r-project.org/web/packages/grf/grf.pdf

39. Cohen J. Statistical Power Analysis for the Behavioral Sciences [Internet]. Routledge; 2013. Available from: https://play.google.com/store/books/details?id=cIJH0lR33bgC

40. Jenni K, Loewenstein G. Explaining the Identifiable Victim Effect. J Risk Uncertain [Internet] 1997;14(3):235–57. Available from: https://doi.org/10.1023/A:1007740225484

41. Morgan SE, Movius L, Cody MJ. The Power of Narratives: The Effect of Entertainment Television Organ Donation Storylines on the Attitudes, Knowledge, and Behaviors of Donors and Nondonors. J Commun [Internet] 2009;59(1):135–51. Available from: https://academic.oup.com/joc/article/59/1/135-151/4098364

42. Duke C, Stanik C. Overcoming Lower-Income Patients' Concerns About Trust And Respect From Providers [Internet]. [cited 2020 Nov 18];Available from: "Overcoming Lower-Income Patients' Concerns About Trust And Respect From Providers, " Health Affairs Blog,

43. Lee JL, DeCamp M, Dredze M, Chisolm MS, Berger ZD. What are health-related users tweeting? A qualitative content analysis of health-related users and their messages on twitter. J Med Internet Res [Internet] 2014;16(10):e237. Available from: http://dx.doi.org/10.2196/jmir.3765

44. Choo EK, Ranney ML, Chan TM, et al. Twitter as a tool for communication and knowledge exchange in academic medicine: A guide for skeptics and novices. Med Teach [Internet] 2015;37(5):411–6. Available from: http://dx.doi.org/10.3109/0142159X.2014.993371

45. Jahng MR, Lee N. When Scientists Tweet for Social Changes: Dialogic Communication and Collective Mobilization Strategies by Flint Water Study Scientists on Twitter. Sci Commun [Internet] 2018;40(1):89–108. Available from: https://doi.org/10.1177/1075547017751948

46. Gallegos A. Hospitals Muzzle Doctors and Nurses on PPE, COVID-19 Cases [Internet]. Medscape. 2020 [cited 2020 Mar 28];Available from: https://www.medscape.com/viewarticle/927541

47. Carville O, Court E, Brown KV. Hospitals Tell Doctors They'll Be Fired If They Speak Out About Lack of Gear [Internet]. Bloomberg News. 2020 [cited 2020 Nov 18];Available from: https://www.bloomberg.com/news/articles/2020-03-31/hospitals-

tell-doctors-they-ll-be-fired-if-they-talk-to-press

48. Friedersdorf C. Hospitals Must Let Doctors and Nurses Speak Out. The Atlantic [Internet] 2020 [cited 2020 Nov 18];Available from: https://www.theatlantic.com/ideas/archive/2020/04/why-are-hospitals-censoring-doctors-and-nurses/609766/

49. Stone W. "It"s Like Walking Into Chernobyl,' One Doctor Says Of Her Emergency Room [Internet]. NPR. 2020 [cited 2020 Nov 18];Available from: https://www.npr.org/2020/04/09/830143490/it-s-like-walking-into-chernobyl-one-doctor-says-of-her-emergency-room

50. Clementson DE. How Web Comments Affect Perceptions of Political Interviews and Journalistic Control: Effects of Web News Attributions. Polit Psychol [Internet] 2019;40(4):815–36. Available from: http://doi.wiley.com/10.1111/pops.12560

51. Brewer PR, Habegger M, Harrington R, Hoffman LH, Jones PE, Lambe JL. Interactivity between Candidates and Citizens on a Social Networking Site: Effects on Perceptions and Vote Intentions. Journal of Experimental Political Science [Internet] 2016 [cited 2020 Nov 18];3(1):84–96. Available from: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/AFA9A266B3BF25D595C011F8A45E0FE1/S2052263014000293a.pdf/div-class-title-interactivity-between-candidates-and-citizens-on-a-social-networking-site-effects-on-perceptions-and-vote-intentions-div.pdf

52. Solnick RE, Peyton K, Kraft-Todd G, Safdar B. Effect of Physician Gender and Race on Simulated Patients' Ratings and Confidence in Their Physicians: A Randomized Trial. JAMA Netw Open [Internet] 2020;3(2):e1920511. Available from: http://dx.doi.org/10.1001/jamanetworkopen.2019.20511

53. Chernozhukov V, Demirer M, Duflo E, Fernández-Val I. Generic machine learning inference on heterogenous treatment effects in randomized experiments. 2018 [cited 2020 Jun 8];Available from: https://open.bu.edu/handle/2144/31469

Figures Legends and Tables

Figure 1. Simulated Twitter Messages for COVID-19 Public Health Messaging
Table 1. Participant Demographics and Baseline Characteristics
Figure 2. Estimated Treatment Effects on Primary Outcomes by Treatment Arm Compared to the Federal, Impersonal Condition
Figure 3. Casual Forest Assessment of Treatment Effect Heterogeneity on Perceived Message Effectiveness by Participant Characteristics

Figure 1. Simulated Twitter Messages for COVID-19 Public Health Messaging

Simulated Twitter posts showing a sample of the federal official/ impersonal treatment arm on the left, and the physician/ personal arm on the right. The text was copied in larger font on the online survey. Two additional posts were created with the texts reversed.

Table 1. Participant Demographics and Baseline Characteristics

	No.(%) of participants by treatment arm				
	Federal Impersonal	Federal Personal	Physician Impersonal	Physician Personal	Overall
	(n=499)	(n=505)	(n=505)	(n=498)	2007
Patient Demographics					
Female	246 (49.3)	247 (48.9)	271 (53.7)	267 (53.6)	1034 (51.4)
Age Group					
18-24	59 (12.4)	70 (14.3)	67 (13.8)	61 (12.7)	257 (13.3)
25-44	187 (39.3)	163 (33.4)	189 (38.9)	178 (36.9)	720 (37.2)
45-64	148 (31.1)	178 (36.5)	152 (31.3)	157 (32.6)	635 (32.8)
65+	82 (17.2)	77 (15.8)	78 (16.0)	86 (17.8)	323 (16.7)
Region (%)					
Midwest	90 (18.0)	107 (21.2)	94 (18.6)	96 (19.3)	388 (19.3)
Northeast	100 (20.0)	115 (22.8)	96 (19.0)	103 (20.7)	414 (20.6)

South	189 (37.9)	184 (36.4)	209 (41.4)	189 (38.0)	772 (38.4)
West	120 (24.0)	99 (19.6)	106 (21.0)	110 (22.1)	436 (21.7)
Race/Ethnicity					
American Indian or Alaskan Native	5 (1.0)	4 (0.8)	4 (0.8)	3 (0.6)	16 (0.8)
Asian	25 (5.0)	27 (5.3)	30 (5.9)	27 (5.4)	110 (5.5)
Black	53 (10.6)	51 (10.1)	59 (11.7)	51 (10.2)	214 (10.6)
Hispanic	57 (11.4)	60 (11.9)	62 (12.3)	55 (11.0)	234 (11.6)
Other	15 (3.0)	16 (3.2)	18 (3.6)	12 (2.4)	61 (3.0)
White	344 (68.9)	347 (68.7)	332 (65.7)	350 (70.3)	1375 (68.4)
Education					
College Grad	291 (58.6)	261 (51.9)	270 (53.6)	299 (60.0)	1122 (56.0)
High School Grad	107 (21.5)	123 (24.5)	115 (22.8)	84 (16.9)	430 (21.4)
No Diploma	12 (2.4)	12 (2.4)	13 (2.6)	10 (2.0)	47 (2.3)
Some College	87 (17.5)	107 (21.3)	106 (21.0)	105 (21.1)	406 (20.2)
Income					
Missing	14 (2.8)	15 (3.0)	21 (4.2)	13 (2.6)	63 (3.1)
<25k	134 (26.9)	117 (23.2)	140 (27.7)	106 (21.3)	498 (24.8)
>99k	108 (21.6)	97 (19.2)	82 (16.2)	114 (22.9)	401 (20.0)
25k-49k	110 (22.0)	118 (23.4)	130 (25.7)	102 (20.5)	461 (22.9)
50k-74k	69 (13.8)	95 (18.8)	83 (16.4)	95 (19.1)	343 (17.1)
75k-99k	64 (12.8)	63 (12.5)	49 (9.7)	68 (13.7)	244 (12.1)
Marital Status					

Married	227 (45.5)	233 (46.1)	233 (46.1)	245 (49.2)	938 (46.7)
Other	130 (26.1)	127 (25.1)	134 (26.5)	121 (24.3)	512 (25.5)
Single	142 (28.5)	145 (28.7)	138 (27.3)	132 (26.5)	557 (27.8)
Health Status					
Missing	6 (1.2)	6 (1.2)	7 (1.4)	6 (1.2)	28 (1.4)
Excellent	67 (13.4)	64 (12.7)	62 (12.3)	75 (15.1)	268 (13.3)
Fair	78 (15.6)	66 (13.1)	76 (15.0)	69 (13.9)	289 (14.4)
Good	189 (37.9)	202 (40.0)	182 (36.0)	191 (38.4)	764 (38.0)
Poor	16 (3.2)	9 (1.8)	15 (3.0)	13 (2.6)	53 (2.6)
Very good	143 (28.7)	158 (31.3)	163 (32.3)	144 (28.9)	608 (30.2)
Baseline Characteristics					
News Frequency					
Frequently	140 (28.1)	156 (30.9)	154 (30.5)	145 (29.1)	595 (29.6)
Other	97 (19.4)	93 (18.4)	92 (18.2)	80 (16.1)	362 (18.0)
Very frequently	262 (52.5)	256 (50.7)	259 (51.3)	273 (54.8)	1050 (52.3)
Prioritize public health over economy	394 (79.1)	396 (78.9)	414 (82.8)	407 (82.1)	1611 (80.7)
Political Party					
Dem	237 (47.5)	229 (45.3)	229 (45.3)	209 (42.0)	905 (45.0)
Ind	60 (12.0)	62 (12.3)	76 (15.0)	69 (13.9)	268 (13.3)
Rep	202 (40.5)	214 (42.4)	200 (39.6)	220 (44.2)	837 (41.6)
Political Ideology					
Missing	6 (1.2)	9 (1.8)	9 (1.8)	6 (1.2)	33 (1.6)

Conservative	101 (20.2)	99 (19.6)	110 (21.8)	101 (20.3)	411 (20.4)
Liberal	93 (18.6)	91 (18.0)	81 (16.0)	79 (15.9)	344 (17.1)
Moderate	191 (38.3)	193 (38.2)	196 (38.8)	197 (39.6)	777 (38.7)
Very conservative	75 (15.0)	72 (14.3)	79 (15.6)	73 (14.7)	299 (14.9)
Very liberal	33 (6.6)	41 (8.1)	30 (5.9)	42 (8.4)	146 (7.3)
Anxiety Level					
Missing	6 (1.2)	6 (1.2)	6 (1.2)	1 (0.2)	22 (1.1)
Not at all	110 (22.0)	94 (18.6)	116 (23.0)	109 (21.9)	429 (21.3)
More than half the days	91 (18.2)	103 (20.4)	97 (19.2)	98 (19.7)	389 (19.4)
Several days	162 (32.5)	185 (36.6)	166 (32.9)	172 (34.5)	685 (34.1)
Nearly every day	130 (26.1)	117 (23.2)	120 (23.8)	118 (23.7)	485 (24.1)

Figure 2. Estimated Treatment Effects on Primary Outcomes by Treatment Arm Compared to the Federal, Impersonal Condition

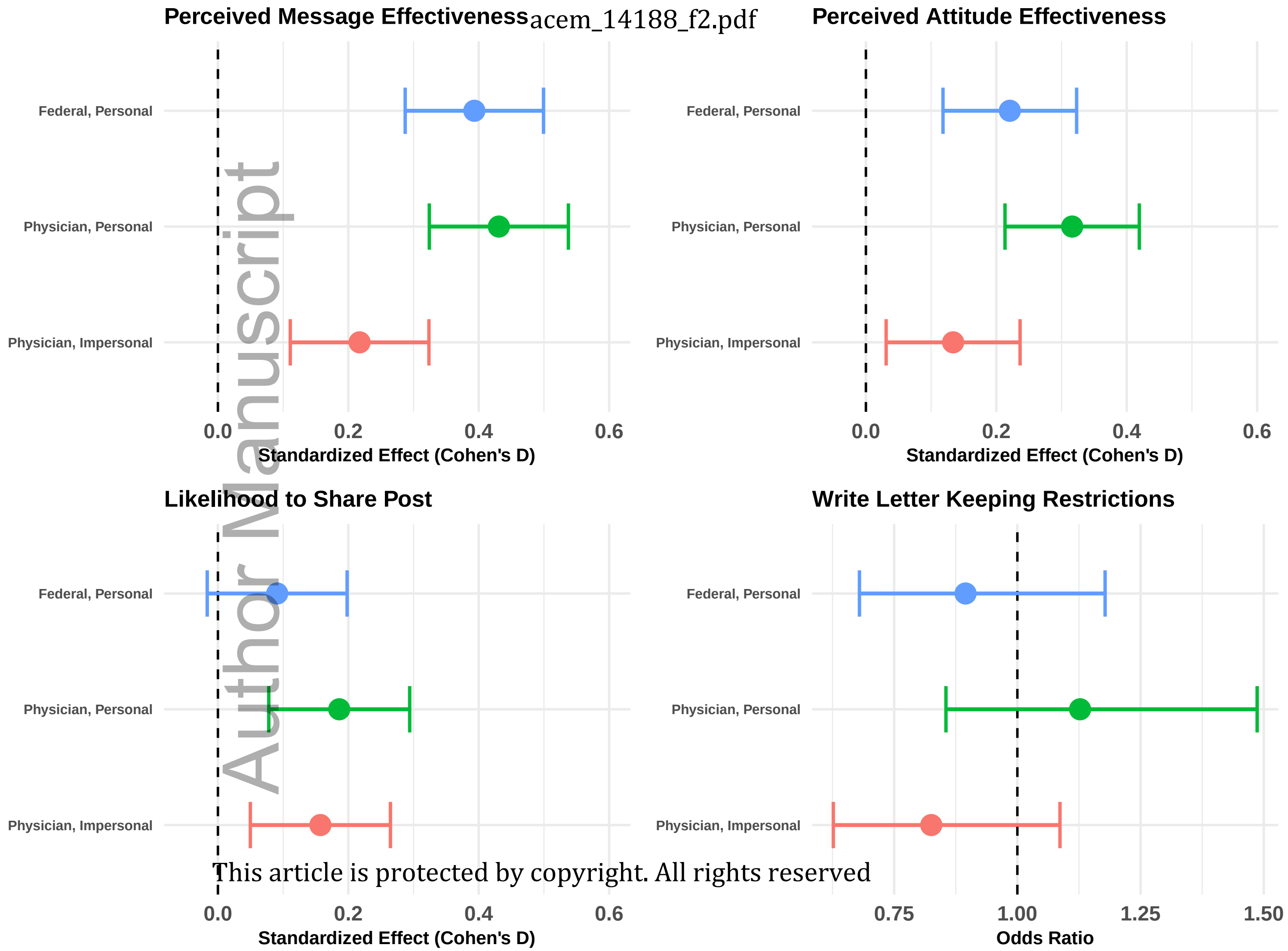
*Covariate-adjusted treatment effects from ordinary least squares regression with reference being the control group, federal impersonal message. Estimates are standardized using Cohen's D, which scales outcomes by the pooled standard deviation. A Cohen's D of 0.2 is considered a small effect and 0.5 a medium effect.[39] (**eTable 6** for tabular form). Points are bounded by 95% CIs. Regression adjusted by covariates: race/ ethnicity, marital (married, single, other), party, gender, anxiety about COVID-19, news frequency (very frequent, frequent, other), and economy versus public health trade-off.*

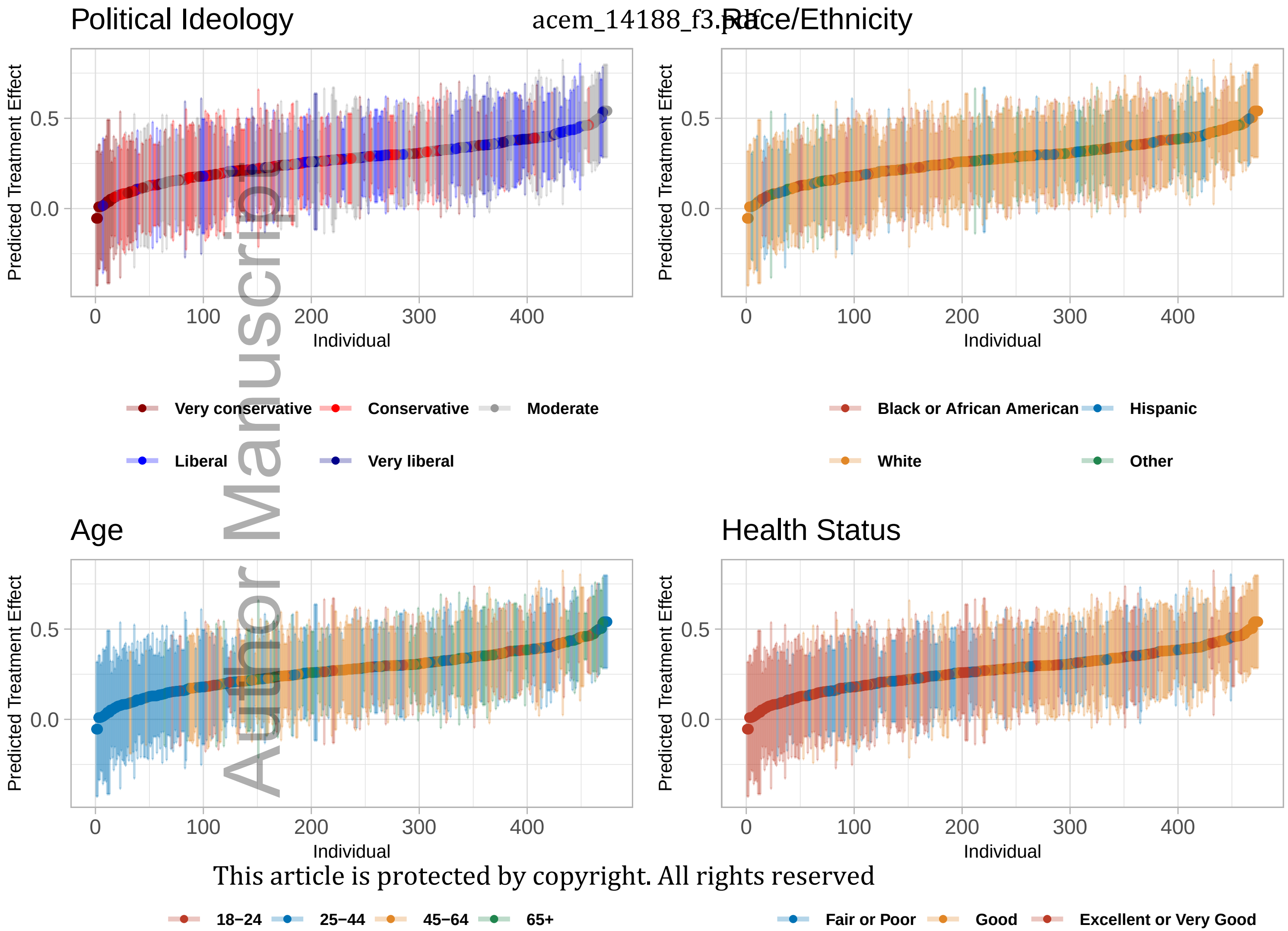
Figure 3. Casual Forest Assessment of Treatment Effect Heterogeneity on Perceived Message Effectiveness by Participant Characteristics

Treatment effect heterogeneity shown for perceived messaging effect outcome, ordered by predicted treatment effect size in Cohen's D standardized units. A Cohen's D of 0.2 is considered a small effect and 0.5 a medium effect.[39] *Omnibus test for heterogeneity*[53] *found no significant heterogeneity in the effect (p-value 0.26). Political ideology and age selected due to highest relative variable importance, though not statistically significant. Race/ethnicity and health status selected due to hypothesized importance, though visually and statistically no heterogeneity demonstrated.*

acem_14188_f1a.tif

acem_14188_f1b.tif


Perceived Message Effectiveness
Federal, Personal
Physician, Personal
Physician, Impersonal
0.0
0.2
0.4
0.6
Standardized Effect (Cohen's D)
Perceived Attitude Effectiveness
Federal, Personal
Physician, Personal
Physician, Impersonal
0.0
0.2
0.4
0.6
Standardized Effect (Cohen's D)
Likelihood to Share Post
Federal, Personal
Physician, Personal
Physician, Impersonal
0.0
0.2
0.4
0.6
Standardized Effect (Cohen's D)
Write Letter Keeping Restrictions
Federal, Personal
Physician, Personal
Physician, Impersonal
0.75
1.00
1.25
1.50
Odds Ratio


Political Ideology
Predicted Treatment Effect
0.5
0.0
0
100
200
300
400
Individual
Very conservative
Conservative
Moderate
Liberal
Very liberal
Race/Ethnicity
Predicted Treatment Effect
Individual
Black or African American
Hispanic
White
Other
Age
Predicted Treatment Effect
Individual
18−24
25−44
45−64
65+
Health Status
Predicted Treatment Effect
Individual
Fair or Poor
Good
Excellent or Very Good