What predicts people's adherence to COVID-19 misinformation? A retrospective study using a nationwide online survey among adults residing in the United States

Sooyoung Kim
New York University School of Global Public Health
Ariadna Capasso
New York University School of Global Public Health
Shahmir H. Ali
New York University School of Global Public Health
Tyler Headley
New York University School of Global Public Health
Ralph J. DiClemente
New York University School of Global Public Health
Yesim Tozan (tozan@nyu.edu)
New York University School of Global Public Health

Research Article

Keywords: COVID-19, misinformation, infodemic, LASSO

Posted Date: April 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1459357/v1

Abstract

Background: Tackling infodemics with flooding misinformation is key to managing the COVID-19 pandemic. Yet only a few studies have attempted to understand the characteristics of the people who believe in and adhere to misinformation.

Methods: Data was used from an online survey that was administered in April 2020 to 6,518 English-speaking adult participants in the United States. We created binary variables to represent four misinformation categories related to COVID-19: general COVID-19-related, vaccine/anti-vaccine, COVID-19 as an act of bioterrorism, and mode of transmission. Using binary logistic regression and the LASSO regularization, we then identified the important predictors of adherence to each type of misinformation. Nested vector bootstrapping approach was used to estimate the standard error of the LASSO coefficients.

Results: About 30% of our sample reported adhering to at least one type of COVID-19-related misinformation. Adherence to one type of misinformation was not strongly associated with adherence to other types. We also identified 58 demographic and socioeconomic factors that predicted people's susceptibility to at least one type of COVID-19 misinformation. Different groups, characterized by distinct sets of predictors, were susceptible to different types of misinformation. There were 25 predictors for general COVID-19 misinformation, 42 for COVID-19 vaccine, 36 for COVID-19 as an act of bioterrorism, and 27 for mode of COVID-transmission.

Conclusion: Our findings confirm the existence of groups with unique characteristics that adheres to different types of COVID-19 misinformation. Findings are readily applicable by policymakers to inform careful targeting of misinformation mitigation strategies.

Background

More than two years have passed since the discovery of SARS-CoV-2, yet the coronavirus disease (COVID-19) pandemic persists.[1,2] Public policies developed and adapted over time in an effort to control the pandemic have faced insufficient engagement and compliance[3], albeit at different levels across the globe. The primary impediments to the enactment of and compliance with best practice public health measures included the uncertain trajectory and the visible social and economic costs of the pandemic, and infodemics. Defined as "too much information including false or misleading information in digital and physical environments during a disease outbreak,"[4] the COVID-19 infodemic polarized opinions and affected compliance with public health measures.[5] Specifically, the proliferation of pandemic-related misinformation have led to the adoption of conspiracy theories and often negatively affected health-related decision-making.[6, 7] For example, a quasi-experimental study conducted in the United States (US) and the United Kingdom (UK) found that people who were exposed to misinformation had a lower intention of getting vaccinated against COVID-19; further, the proportion of vaccine-hesitant people had grown in tandem with the spread of misinformation.[8] Moreover, misinformation impacted

sociodemographic groups differently,[8] highlighting the importance of targeted communication to achieving herd immunity.

The problem of widespread misinformation is not new. With user-generated content inundating online platforms like social media, effectively countering misinformation has long been a challenge in the field of public health.[9–12] One demonstrated method to thwart misinformation is through active and strategic responses based on demonstrating misinformation's falsehood[6, 7] and presenting the correct information through targeted dissemination.[7, 13–15] Understanding how groups respond to misinformation is therefore critical to creating, targeting, and executing a counter-misinformation strategy.

However, studies on COVID-19 misinformation have primarily focused on profiling the types and sources of misinformation,[16–19] detecting misinformation using machine learning algorithms,[20–24] or exploring the behavior-related consequences of misinformation.[16, 25–29] Only a few have attempted to understand the characteristics of the people or communities who believe in and adhere to COVID-19 misinformation.[30, 31] Roozenbeek and colleagues used a cross-sectional survey from five countries (Ireland, the US, Mexico, Spain, and the UK) to identify the predictors of susceptibility to misinformation.[30] They regressed an average susceptibility score on a pre-selected set of predictors and found that susceptibility to misinformation was negatively associated with compliance with COVID-19 public health guidance, including willingness to get vaccinated. Lobato *et al.* conducted an exploratory canonical correlation analysis to identify individual characteristics associated with willingness to share misinformation.[31] The authors found that certain aspects of political beliefs predicted tendencies to disseminate misinformation.

This study expands on prior studies on COVID-19 misinformation in two important ways. First, we employed a novel model selection approach to widen the scope of potential predictors rather than narrowing the scope to a pre-selected subset. Second, we explored people's adherence to different types of misinformation, informed by the published literature,[16, 18, 30, 32] rather than aggregating misinformation into a single index. The primary aim of the study was to provide insights into who adheres to COVID-19 misinformation so as to inform the design and targeting of misinformation mitigation strategies.

Methods

Data

This study is a secondary analysis of data from an online survey conducted in April 2020. The primary aim of the survey was to collect and analyze data on COVID-19-related knowledge, beliefs, and behaviors among the US adult population during the early days of the pandemic.[33] While the details of the survey design and administration methods are described elsewhere,[33] in short, the questionnaire was developed based on the Health Belief Model[34] and validated scales from the literature.[33] Participants were recruited using a convenience sampling approach on Facebook and its affiliated platforms through social media

advertisement campaigns. The survey was completed by 6,518 voluntary and eligible participants. Eligibility criteria included being an English-speaking adult (aged 18 years and older) who was physically residing in the US. The analysis only included participants who expressed informed consent and provided a complete response to a meaningful subset of the survey questionnaire. We checked the pattern of missingness in the data, specifically whether data are missing completely at random (MCAR) or missing at random (MAR).[35, 36] The study protocol was reviewed and deemed exempt by New York University's Institutional Review Board.

Adherence to misinformation

We derived four misinformation variables—general, bioterrorism, anti-vaccine, and transmission mode—from our survey. First, we created a variable on adherence to generalized misinformation by identifying respondents whose COVID-19 knowledge scores were in the bottom quartile[37] and who responded "yes" to the statement "I believe the information I get about Coronavirus is accurate." Next, we derived variables for adherence to bioterror, anti-vaccine, and transmission mode misinformation. Prior studies showed that COVID-19 misinformation clustered into distinct thematic categories and that different "dubious beliefs" about COVID-19 attracted distinct groups of people.[16, 18, 32] While there is no single agreed-upon approach to this categorization, the most common categories of misinformation include the modes of transmission; miracle cures or treatments; anti-vaccine; political conspiracy theories; racism; and bioterrorism.[16–19, 30, 32]

Using the variables in our data set we created binary variables for the three dubious beliefs. First, we classified participants as adhering to misinformation on bioterrorism if they responded "strongly agree" or "agree" to the statement "I think that Coronavirus was released as an act of bioterrorism." Second, we classified participants as adhering to misinformation related to a COVID-19 vaccine if they responded "not likely" to the question "how likely would you be to get a Coronavirus vaccine if it was recommended by: doctor/medical provider?" We note that though no vaccine had been released at the time of this survey, there was already a significant volume of misinformation about possible COVID-19 vaccines, such as the conspiracy theory that the vaccines would include a geolocation-tracking microchip; thus, reticence to get a hypothetical vaccine that was hypothetically endorsed by respondents' medical providers was categorized as adherence to misinformation. Third, we classified participants as adhering to misinformation on the mode of COVID-19 transmission if they: 1) answered "no" to "practicing social distancing" and "wearing a face mask or covering when they leave home," and 2) responded "strongly disagree" or "disagree" to the statement "if I were ORDERED to quarantine myself due to Coronavirus, I would do so."

Potential predictors of adherence to misinformation

We included 66 variables from the survey as potential predictors of people's adherence to misinformation. These variables depicted participants' sociodemographic characteristics, including, but not limited to, age, sex, race/ethnicity, highest educational attainment, annual household income, marital status, residence area, political affiliation, COVID-19-related knowledge levels, information-seeking patterns, and

beliefs and perceptions about the COVID-19 disease. A complete list of the variables, their definitions, and participants' responses are summarized in the Supplementary Material (Supplementary Material 1, Table S1).

Data analysis

We used binary logistic regressions and the LASSO (Least Absolute Shrinkage and Selection Operator) regularization to select important predictors of adherence to misinformation among the initial set of 66 variables (p = 66). The equation below illustrates the logistic regression of the outcome variable of misinformation adherence (Y) using the set of *p* predictors ($X_1, X_2, \ldots, X_p$).

$$\log\left(\frac{Pr(Y=1)}{1-Pr(Y=1)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

LASSO regularization reduces the high dimensionality of the data. It is a variable selection method that has been increasingly used in place of the traditional stepwise selection approach (i.e., backward selection, forward selection).[35] It has been shown to improve the model fit by avoiding stepwise selection's path-dependency and reducing overfitting issues by using a cross-validation approach.[35] In brief, the LASSO method enables the selection of a model with the best fitting subset of explanatory variables by introducing the penalty term $\lambda \sum_{j=1}^{p} \left|\beta_j\right|$ into the regression equation. The method estimates the regression coefficients (β_j) by minimizing the sum of squared residuals ($\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2$) while shrinking some of the coefficient estimates to zero when the tuning parameter λ is sufficiently large. As a result, the LASSO logistic regression yields a sparse model with only a subset of variables.

$$\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 + \lambda \sum_{j=1}^{p} \left|\beta_j\right|$$

We used the glmnet package in R[38] to conduct the LASSO logistic regression. We used 10-fold cross-validation to identify the optimal value of λ. We then followed the vector bootstrapping approach proposed by Laurin et al.[39] to estimate the standard error (SE) and the 95% confidence interval (CI) of the LASSO coefficients ($\hat{\beta}_j$). We took this additional step because, first, we wanted to account for the low variable selection precision (VSP), the percent of true important predictors among the model-selected predictors, following the LASSO approach.[40–42] Second, we wanted to improve the interpretability of the model by constructing the 95% CI for the LASSO estimate. By doing so, the results can be interpreted similarly to the conventional frequentist framework.[39] Specifically, we used the nested cross-validated selection methods for λ (described as Method 3 in Laurin et al., 2016), which leads to larger SEs than the fixed λ bootstrapping (Method 2).[39] Using Monte Carlo simulation, we estimated the coefficients and

calculated the 95% CI of the coefficients based on an approximate inverted z-test ($\hat{\beta}_j \pm z_{\alpha/2} * SE * (\hat{\beta}_j)$), where $z_{\alpha/2}$ is the $\frac{\alpha}{2}$ quantile of a standard normal distribution and $\alpha = 0.05$ for the 95% CI. We kept the variables that had non-zero coefficients with their 95% confidence interval not crossing the value zero. Thus, the final set of variables retained in the model further improve the VSP. The larger estimated SE from the nested cross-validation (CV) approach for λ also made the CI-based selection of variables more conservative than the fixed λ bootstrapping.[39]

Results

Descriptive statistics

Table 1 summarizes the descriptive statistics of the survey participants. Of the 6,518 survey participants, only 2,793 (42.9%) provided a complete response to the variables used in the analysis. Upon checking the missingness of the data, while most explanatory variables were missing completely at random (MCAR), some variables, namely highest educational attainment, annual household income, employment status, and type of residence, showed a missingness pattern at random (MAR) (Supplementary Material, Figures S1.1 and S1.2). The distribution of several sociodemographic characteristics such as sex, race, highest educational attainment, and political affiliation to Democratic Party was similar between the overall sample and the regression sample (the data subset which included only complete responses). However, the distribution of age group, marital status, the number of children and people in a household, employment status, annual household income, political affiliation to Republican Party, and geographic region were significantly different between two samples.

Table 1
Descriptive statistics of the COVID-19 survey responses in April 2020 for the total sample (N = 6,518) and the regression sample with complete data only (N = 2,793)

	Total sample (N = 6,518)	**Regression Sample (N = 2,793)**	**p-value**
Sex			**0.951**
Female	3717 (57.0%)	1610 (57.6%)	
Male	2738 (42.0%)	1183 (42.4%)	
Missing	63 (1.0%)		
Age group			**< 0.001**
18–29 years old	343 (5.3%)	120 (4.3%)	
30–39 years old	735 (11.3%)	372 (13.3%)	
40–49 years old	997 (15.3%)	495 (17.7%)	
50–59 years old	1814 (27.8%)	863 (30.9%)	
60–69 years old	1967 (30.2%)	755 (27.0%)	
70–79 years old	605 (9.3%)	179 (6.4%)	
80 + years old	57 (0.9%)	9 (0.3%)	
Race			**0.051**
White, Non-Hispanic	6012 (92.2%)	2634 (94.3%)	
Hispanic/Latinx	169 (2.6%)	52 (1.9%)	
Interracial, Mixed race, or Other	190 (2.9%)	63 (2.3%)	
Asian/Pacific Islander	50 (0.8%)	15 (0.5%)	
Black, Non-Hispanic	53(0.8%)	12 (0.4%)	
Native American or American Indian	44 (0.7%)	17 (0.6%)	
Currently married			**< 0.001**
No	1475 (22.6%)	492 (17.6%)	
Yes	3585 (55.0%)	2301 (82.4%)	
Missing	1458 (22.4%)		
Children under 18 in the household			**< 0.001**
No	4253 (65.3%)	1893 (67.8%)	

	Total sample (N = 6,518)	Regression Sample (N = 2,793)	p-value
Yes	1477 (22.7%)	900 (32.2%)	
Missing	788 (12.1%)		
Number of people in the household			**0.015**
Mean (SD)	3.16 (1.70)	2.84 (1.26)	
Employment status			**< 0.001**
Employed	2845 (43.6%)	1832 (65.6%)	
Student/Unpaid work	280 (4.3%)	140 (5.0%)	
Not working/Unemployed	635 (9.7%)	325 (11.6%)	
Retired	1300 (19.9%)	496 (17.8%)	
Missing	1458 (22.4%)		
Highest educational attainment			**0.050**
High school degree / GED or less	516 (7.9%)	264 (9.5%)	
Some college / Associate's degree	1720 (26.4%)	944 (33.8%)	
Bachelor's degree or higher	2792 (42.8%)	1585 (56.7%)	
Missing	1490 (22.9%)		
Annual household income			**< 0.001**
Less than $30,000	580 (8.9%)	233 (8.3%)	
$30,000 to less than $50,000	671 (10.3%)	378 (13.5%)	
$50,000 to less than $75,000	767 (11.8%)	477 (17.1%)	
$75,000 to less than $100,000	900 (13.8%)	614 (22.0%)	
$100,000 or more	1419 (21.8%)	1091 (39.1%)	
Missing	2181 (33.5%)		
Democrat (political affiliation)			**0.675**
No	3103 (47.6%)	1716 (61.4%)	
Yes	1925 (29.5%)	1077 (38.6%)	
Missing	1490 (22.9%)		
Republican (political affiliation)			**< 0.001**

	Total sample (N = 6,518)	Regression Sample (N = 2,793)	p-value
No	3806 (58.4%)	2043 (73.1%)	
Yes	1222 (18.7%)	750 (26.9%)	
Missing	1490 (22.9%)		
Region of residence			**0.042**
Northeast	1379 (21.2%)	772 (27.6%)	
Midwest	1308 (20.1%)	756 (27.1%)	
South	1379 (21.2%)	746 (26.7%)	
West	994 (15.3%)	519 (18.6%)	
Missing	1458 (22.4%)		
Type of residence			**0.009**
Suburban	2697 (41.4%)	1538 (55.1%)	
Urban	770 (11.8%)	395 (14.1%)	
Rural	1593 (24.4%)	860 (30.8%)	
Missing	1458 (22.4%)		

Overall, 31.4% (n = 2048) of the total respondents and 35.2% (n = 982) of the respondents included in the regression sample adhered to at least one type of misinformation. In the overall sample, 23.9% (n = 794) adhered to bioterrorism misinformation, 12.7% (n = 826) believed in misinformation about a hypothetical COVID-19-vaccine, 4.5% (n = 294) of the respondents adhered to general misinformation, and 1.8% (n = 120) believed misinformation about the mode of COVID-19 transmission. The proportion of people adhering to misinformation was generally similar between the total and the regression sample, as shown in Fig. 1. Interestingly, adherence to one type of misinformation was not strongly associated with adherence to other types of misinformation. While the overall prevalence of adherence to any type of misinformation was estimated to be 35.2% (n = 982) in the regression sample, only 8.8% (n = 246) of the participants adhered to two types of misinformation, 2% (n = 55) adhered to three types of misinformation, and 0.1% (n = 3) adhered to all four types of misinformation (Supplementary Material 2, Figure S2-1). The strongest correlation (coefficient = 0.32) was observed between the adherence to misinformation related to the hypothetical COVID-19 vaccine and bioterrorism, followed by the relationship between adherence to anti-vaccine misinformation and modes of transmission (coefficient = 0.25). Cross-tabulation of adherence to different types of misinformation is provided in the supplementary material (Supplementary Material 2, Table S2-1 ~ 7).

LASSO logistic regressions

Figure 2 summarizes the results of the vector-bootstrapped LASSO logistic regression on the factors associated with adherence to misinformation. A total of 58 factors were significantly associated with adherence to at least one type of misinformation. Among them, 38 factors were positively associated and 38 were negatively associated with endorsement of at least one type of misinformation. Only two predictors, never searching COVID-19 information online and not using mainstream media as COVID-19 information source, was associated with significantly increased odds of adhering to all four types of misinformation. Additionally, respondents' highest educational attainment being high school or less or some college/associate's degree predicted adherence to three types of misinformation—general, anti-vaccine, and bioterrorism misinformation. Being Native American or American Indian, or of mixed race, male, Republican, a resident of the South or earning annual household income less than $30,000 was a common predictor for two different types of misinformation, as was using Fox News, a religious leader, or social media as a COVID-19 information source. Conversely, using a newspaper or the government's official communication as a source of COVID-19-related information or having a health insurance coverage was associated with significantly lower odds of adhering to all types of misinformation. Higher COVID-19-related knowledge or using TV as COVID-19 information source similarly predicted significantly lower odds of adhering to misinformation related to the hypothetical COVID-19 vaccine, bioterrorism, and modes of transmission.

Interestingly, 18 predictors worked in opposite directions for different types of misinformation. For example, respondents' age being 80 and above was a predictor for higher odds of adhering to general misinformation, but was associated with lower odds of adhering to bioterrorism or anti-vaccine misinformation. While those who used a mental health service due to COVID-19 reported higher odds of believing general misinformation, the use of a mental health service was associated with decreased odds of believing anti-vaccine or transmission mode misinformation. Similarly, people with high levels of anxiety, who are retired, or who reported to be a healthcare worker had higher odds of believing bioterrorism misinformation and decreased odds of believing in anti-vaccine misinformation.

Figure 3 presents the predictors of each misinformation type in descending order by effect size. Respondents had higher odds of adhering to general misinformation if they were Black, non-Hispanic or Native American/American Indian, aged 80 years and above, male, seeking mental health services for COVID-19, having highest educational attainment of high school or less degree, earning less than $50,000 annual household income, using Fox News as a COVID-19 information source, or never seeking COVID-19 information. Adherence to anti-vaccine misinformation was most strongly associated with being mixed race or Native American/American Indian, never seeking COVID-19-related information, or lower educational attainment of some college or associate's degree and below. Adherence to bioterrorism misinformation was strongly and significantly associated with many predictors, including being politically affiliated with the Republican party, having low educational attainment of college or associate's degree or less, food insecure, and a user of Fox News, social media, or a religious leader as the primary COVID-19 information source. Finally, adherence to misinformation on the mode of transmission was

most strongly associated with never seeking COVID-19-related information, being male, having moved residence due to COVID-19, or reporting a higher level of loneliness. Summarized adjusted odds ratios and the 95% bootstrap CIs are available in the Supplementary Material (Supplementary Material 3, Table S3-1 ~ 4).

Discussion

Countering infodemics with targeted, factual information is crucial for ending the COVID-19 pandemic.[4, 7] Understanding what factors have played a role in people's adherence to COVID-19 misinformation is critical to enabling policymakers to craft strategic communications to manage the COVID-19 infodemic and may provide insights on how to tackle the future infodemics related to novel infectious disease threats. Our study attempted to identify the factors associated with adherence to certain types of COVID-19 misinformation among US adults in April 2020 and showed that misinformation started affecting the general public from the early phase of the pandemic. We performed our analysis on four types of misinformation: general, anti-vaccine, bioterrorism, and transmission modes. Our use of LASSO regressions allowed us to identify and select significant predictors from a broad pool of potential factors while simultaneously reducing selection bias, a marked improvement on the traditional approach of using predetermined small subset of predictors. Using bootstrapping, we further refined predictor selection and quantified our estimates' standard error, which increased our confidence in our results.

First, we found that more than 30% of our sample of US adults on social media reported adhering to at least one type of COVID-19-related misinformation in early 2020. Compared to later estimates,[43, 44] our findings suggest that the prevalence of misinformation grew in tandem with the infodemic, as has been noted by health authorities and researchers.[7, 8] This suggests that it is important to promptly address the spread of misinformation, lest an infodemic grow out of control. Second, we found that particular demographic and socioeconomic factors predicted respondents' susceptibility to COVID-19 misinformation. Of the 66 variables included in our analysis, 58 were significantly associated with increased or decreased odds of adhering to specific types of misinformation about COVID-19. Many of these variables were characteristics that are readily available and routinely collected as part of other national surveys, which suggests that policymakers could develop and leverage cost-effective predictive models using existing datasets to identify specific communities and individuals more likely to adhere to misinformation.

Third, we found that different audiences were susceptible to different types of misinformation. Prior research on COVID-19 misinformation tended to aggregate all types of misinformation into a unified index despite the weak correlation between the types of misinformation.[30] This method overlooked key differences and made it difficult to identify differences between sociodemographic groups' adherence to misinformation, which in turn led to policymakers treating everyone who adheres to any COVID-19 misinformation as one target group for interventions. As previously noted, anti-misinformation communication strategies need to be targeted to specific subgroups to be effective.[15, 45] Our findings

confirmed that there are clear differences in subgroups' adherence to misinformation, and should be used to inform strategies that effectively engage those groups by understanding their existing beliefs and motivations, and that address the structural and economic factors that facilitate or promote adherence to misinformation.[15]

Our study had several limitations. First, we used nonprobability convenience sampling via social media platforms affiliated with Facebook to collect our survey data. Because of this approach, our sample may not be representative of the US adult population, despite our sample's large sample size.[33] While our sample is balanced across geography, age groups, and other sociodemographic characteristics, we acknowledge the under-representation of certain subpopulations that might be particularly vulnerable to misinformation. For example, our choice of sampling platform systemically excluded people without access to the internet or social media platforms. Given time constraints and the impracticality of face-to-face recruitment due to the COVID-19 pandemic, we chose the social media platforms affiliated with Facebook as a recruitment and dissemination platform to maximize our reach to the general US population; 70% of the US population are estimated to have Facebook accounts, and among those with accounts, 75% use Facebook daily.[46] Foreign-born adults with limited English-speaking skills, comprising over 40 million adults,[47] would additionally have been excluded from our sample. As a result, the participants in our study were overwhelmingly non-Latinx white despite our concerted effort to oversample potentially under-represented sociodemographic groups.[33] Thus, given these sampling limitations, the findings from our study cannot be generalized to the US population. In particular, those under-represented subpopulations may likely provide further insights into different subgroups who adhere to misinformation. Several studies have highlighted immigrants' elevated risk of exposure to misinformation and difficulty in accessing needed information and resources during the COVID-19 pandemic.[48–50] Further research focusing on this under-represented community is therefore warranted.

Second, our regression analysis was conducted on a subset of the sample that only contains complete responses with no missing data. Since most missing data were MCAR, we did not perform any imputations. Differential missingness of certain responses which were MAR and whose missingness is associated with the outcome variables,[36, 51] however, might have caused some bias. Despite this limitation, we believe our findings still provide novel and significant insights on specific groups. Further, we believe that the benefits of our methodological approach—namely the LASSO regressions, which require complete data—outweigh the costs of subsetting our dataset.

Third, we note that our misinformation categories, which were formulated in the nascent stages of the pandemic, were not necessarily the categories of misinformation that ultimately played the most significant role in individuals' adherence to—or rejection of—public health guidelines. For instance, "bioterrorism" ultimately had less bearing on the public than other strains of misinformation, and anti-vaccine information proliferated and became more nuanced after the first vaccines were released. Future research should seek to investigate specific strains of misinformation, such as "vaccine chip" misinformation versus "vaccine poison" misinformation. Finally, we note that the number of respondents

who adhered to transmission mode information was very small (N = 56), and that particularly in the US context, an underlying cultural lack of conformity with government mandates may have also contributed to individuals responding that they would not comply with government mandated social distancing. The transmission mode misinformation results should therefore be interpreted with some skepticism.

Conclusions

The proliferation of user-generated content on social media has accelerated and perpetruated the spread of misinformation.[4, 7] This misinformation can play a significant role in misdirecting individuals' decision making and adherence to public health guidelines, which in turn hinders effective control of the COVID-19 pandemic. To effectively counter misinformation, targeted communication to specific subgroups of the population is necessary; our findings will enable policymakers to improve their messaging's efficacy against COVID-19 misinformation. Our study further establishes research best practices in the early stages of an epidemic or pandemic, and demonstrates the use of a novel methodology to pinpoint adherence to specific types of misinformation.

Declarations

Ethics approval and consent to participate: The study protocol was reviewed and deemed exempt by New York University's Institutional Review Board. All survey respondents expressed their informed consent before participating in the survey. All methods used in the study were carried out in accordance with relevant guidelines and regulations.

Consent for publication: Not applicable

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: The study is not funded.

Authors' contributions: SK, YT, AC, RJC conceptualized the study. SK conducted the statistical analysis and drafted the first version of the manuscript. AC, SHA, RJC, and YT was involved in survey design and data collection. All authors contributed to, read and approved the final manuscript.

Acknowledgements: Not applicable

References

1. World Health Organization. Archived: WHO Timeline - COVID-19. In; 2020.

2. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. [Dashboard] 2020 Updated daily [cited 2020 16 October 2020]; Available from: https://covid19.who.int/
3. Organisation for Economic Co-operation and Development. Key policy responses from the OECD. 2021 [cited 2022 Jan 6, 2022]; Available from: https://www.oecd.org/coronavirus/en/policy-responses
4. World Health Organization. Infodemic Management - Infodemiology. 2020 [cited 2020 16 October 2020]; Available from: https://www.who.int/teams/risk-communication/infodemic-management
5. Weible CM, Nohrstedt D, Cairney P, Carter DP, Crow DA, Durnová AP, et al. COVID-19 and the policy sciences: initial reactions and perspectives. Policy Sciences 2020;53(2):225–241.
6. Southwell BG, Niederdeppe J, Cappella JN, Gaysynsky A, Kelley DE, Oh A, et al. Misinformation as a Misunderstood Challenge to Public Health. Am J Prev Med 2019;57(2):282–285.
7. Pan American Health Organization. Understanding the infodemic and misinformation in the fight against COVID-19. 2020.
8. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nature Human Behaviour 2021;5(3):337–348.
9. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. Health communication 2018;33(9):1131–1140.
10. Chou W-YS, Oh A, Klein WM. Addressing health-related misinformation on social media. Jama 2018;320(23):2417–2418.
11. Swire-Thompson B, Lazer D. Public health and online misinformation: Challenges and recommendations. Annual Review of Public Health 2019;41:433–451.
12. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Social science & medicine 2019;240:112552.
13. Zarocostas J. How to fight an infodemic. The Lancet 2020;395(10225).
14. Goodwin R, Wiwattanapantuwong J, Tuicomepee A, Suttiwan P, Watakakosol R. Anxiety and public responses to covid-19: Early data from Thailand. J Psychiatr Res 2020;129:118–121.
15. French J, Deshpande S, Evans W, Obregon R. Key Guidelines in Developing a Pre-Emptive COVID-19 Vaccination Uptake Promotion Strategy. Int J Environ Res Public Health 2020;17(16).
16. Enders AM, Uscinski JE, Klofstad C, Stoler J. The different forms of COVID-19 misinformation and their consequences. The Harvard Kennedy School Misinformation Review 2020.
17. Shahi GK, Dirkson A, Majchrzak TA. An exploratory study of COVID-19 misinformation on Twitter. Online Social Networks and Media 2021;22:100104.
18. Evanega S, Lynas M, Adams J, Smolenyak K, Insights CG. Coronavirus misinformation: quantifying sources and themes in the COVID-19 'infodemic'. JMIR Preprints 2020;19(10):2020.
19. Brennen JS, Simon FM, Howard PN, Nielsen RK. Types, sources, and claims of COVID-19 misinformation: University of Oxford; 2020.

20. Bang Y, Ishii E, Cahyawijaya S, Ji Z, Fung P. Model generalization on COVID-19 fake news detection. arXiv preprint arXiv:2101.03841 2021.
21. Wani A, Joshi I, Khandve S, Wagh V, Joshi R. Evaluating deep learning approaches for covid19 fake news detection. In: International Workshop on Combating On line Ho st ile Posts in Regional Languages dur ing Emerge ncy Si tuation; 2021: Springer; 2021. p. 153–163.
22. Patwa P, Sharma S, Pykl S, Guptha V, Kumari G, Akhtar MS, et al. Fighting an infodemic: Covid-19 fake news dataset. In: International Workshop on Combating On line Ho st ile Posts in Regional Languages dur ing Emerge ncy Si tuation; 2021: Springer; 2021. p. 21–29.
23. Hossain T, Logan IV RL, Ugarte A, Matsubara Y, Young S, Singh S. COVIDLies: Detecting COVID-19 misinformation on social media. 2020.
24. Al-Rakhami MS, Al-Amri AM. Lies kill, facts save: detecting COVID-19 misinformation in twitter. Ieee Access 2020;8:155961–155970.
25. Lee JJ, Kang K-A, Wang MP, Zhao SZ, Wong JYH, O'Connor S, et al. Associations Between COVID-19 Misinformation Exposure and Belief With COVID-19 Knowledge and Preventive Behaviors: Cross-Sectional Online Study. J Med Internet Res 2020;22(11):e22205.
26. Chen E, Chang H, Rao A, Lerman K, Cowan G, Ferrara E. COVID-19 misinformation and the 2020 U.S. presidential election. Harvard Kennedy School Misinformation Review 2021.
27. Bridgman A, Merkley E, Loewen PJ, Owen T, Ruths D, Teichmann L, et al. The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. Harvard Kennedy School Misinformation Review 2020;1(3).
28. Barua Z, Barua S, Aktar S, Kabir N, Li M. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. Progress in Disaster Science 2020;8:100119.
29. Tasnim S, Hossain MM, Mazumder H. Impact of Rumors and Misinformation on COVID-19 in Social Media. J Prev Med Public Health 2020;53(3):171–174.
30. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020;7(10):201199.
31. Lobato EJC, Powell M, Padilla LMK, Holbrook C. Factors Predicting Willingness to Share COVID-19 Misinformation. Front Psychol 2020;11:566108.
32. Charquero-Ballester M, Walter JG, Nissen IA, Bechmann A. Different types of COVID-19 misinformation have different emotional valence on Twitter. Big Data & Society 2021;8(2).
33. Ali SH, Foreman J, Capasso A, Jones AM, Tozan Y, DiClemente RJ. Social media as a recruitment platform for a nationwide online survey of COVID-19 knowledge, beliefs, and practices in the United States: methodology and feasibility analysis. BMC Med Res Methodol 2020;20(1):116.
34. Rosenstock IM. The Health Belief Model and Preventive Health Behavior. Health Education Monographs 1974;2(4):354–386.
35. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. Springer; 2013.

36. Harrison E. Missing data. Available from: https://cran.r-project.org/web/packages/finalfit/vignettes/missing.html
37. Kim S, Capasso A, Cook SH, Ali SH, Jones AM, Foreman J, et al. Impact of COVID-19-related knowledge on protective behaviors: The moderating role of primary sources of information. PloS one 2021;16(11):e0260643.
38. Hastie T, Qian J. Glmnet vignette. Retrieved June 2014;9(2016):1–30.
39. Laurin C, Boomsma D, Lubke G. The use of vector bootstrapping to improve variable selection precision in Lasso models. Stat Appl Genet Mol Biol 2016;15(4):305–20.
40. Devlin B, Roeder K, Wasserman L. Analysis of multilocus models of association. Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society 2003;25(1):36–47.
41. Ayers KL, Cordell HJ. SNP selection in genome-wide and candidate gene studies via penalized logistic regression. Genetic epidemiology 2010;34(8):879–891.
42. He Q, Lin D-Y. A variable selection method for genome-wide association studies. Bioinformatics 2011;27(1):1–8.
43. Willis DE, Andersen JA, Bryant-Moore K, Selig JP, Long CR, Felix HC, et al. COVID-19 vaccine hesitancy: Race/ethnicity, trust, and fear. Clin Transl Sci 2021;14(6):2200–2207.
44. King WC, Rubinstein M, Reinhart A, Mejia R. Time trends, factors associated with, and reasons for COVID-19 vaccine hesitancy in a massive online survey of US adults: January-May 2021. medRxiv 2021:2021.07.20.21260795.
45. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. American journal of health behavior 2003;27(1):S227-S232.
46. Smith A, Anderson M. Social Media Use in 2018. 2018 March 1, 2018 [cited 2020 March 29]; Available from: https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/
47. Gambino CP, Acosta YD, Grieco EM. English-Speaking Ability of the Foreign-Born Population in the United States: 2012. Washinton, D.C.: U.S. Census Bureau; 2014.
48. Bastick Z, Mallet M. Double lockdown: The effects of digital exclusion on undocumented immigrants during the COVID-19 pandemic. Available at SSRN 3883432 2021.
49. Ross J, Diaz CM, Starrels JL. The Disproportionate Burden of COVID-19 for Immigrants in the Bronx, New York. JAMA Intern Med 2020;180(8):1043–1044.
50. Deal A, Hayward SE, Huda M, Knights F, Crawshaw AF, Carter J, et al. Strategies and action points to ensure equitable uptake of COVID-19 vaccinations: A national qualitative interview study to explore the views of undocumented migrants, asylum seekers, and refugees. Journal of Migration and Health 2021;4:100050.
51. Van Buuren S. Flexible imputation of missing data. CRC press; 2018.

Figures

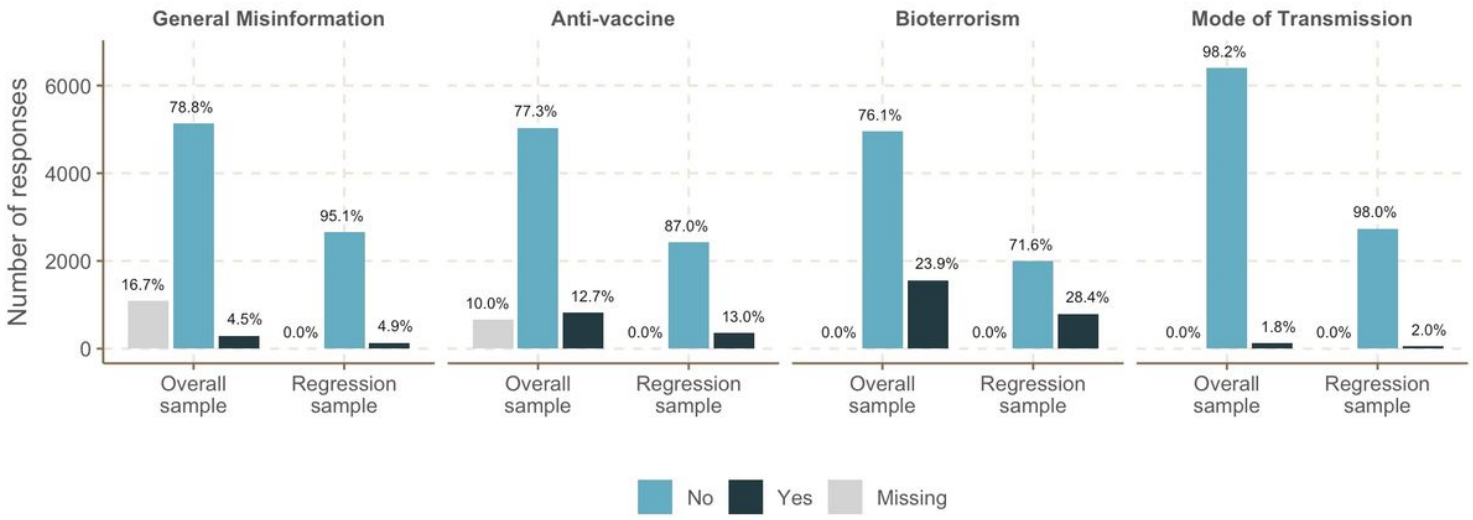

Figure 1

Distribution of survey participants adhering to different types of misinformation in the total sample (N=6,518) and the regression sample (N=2,793)

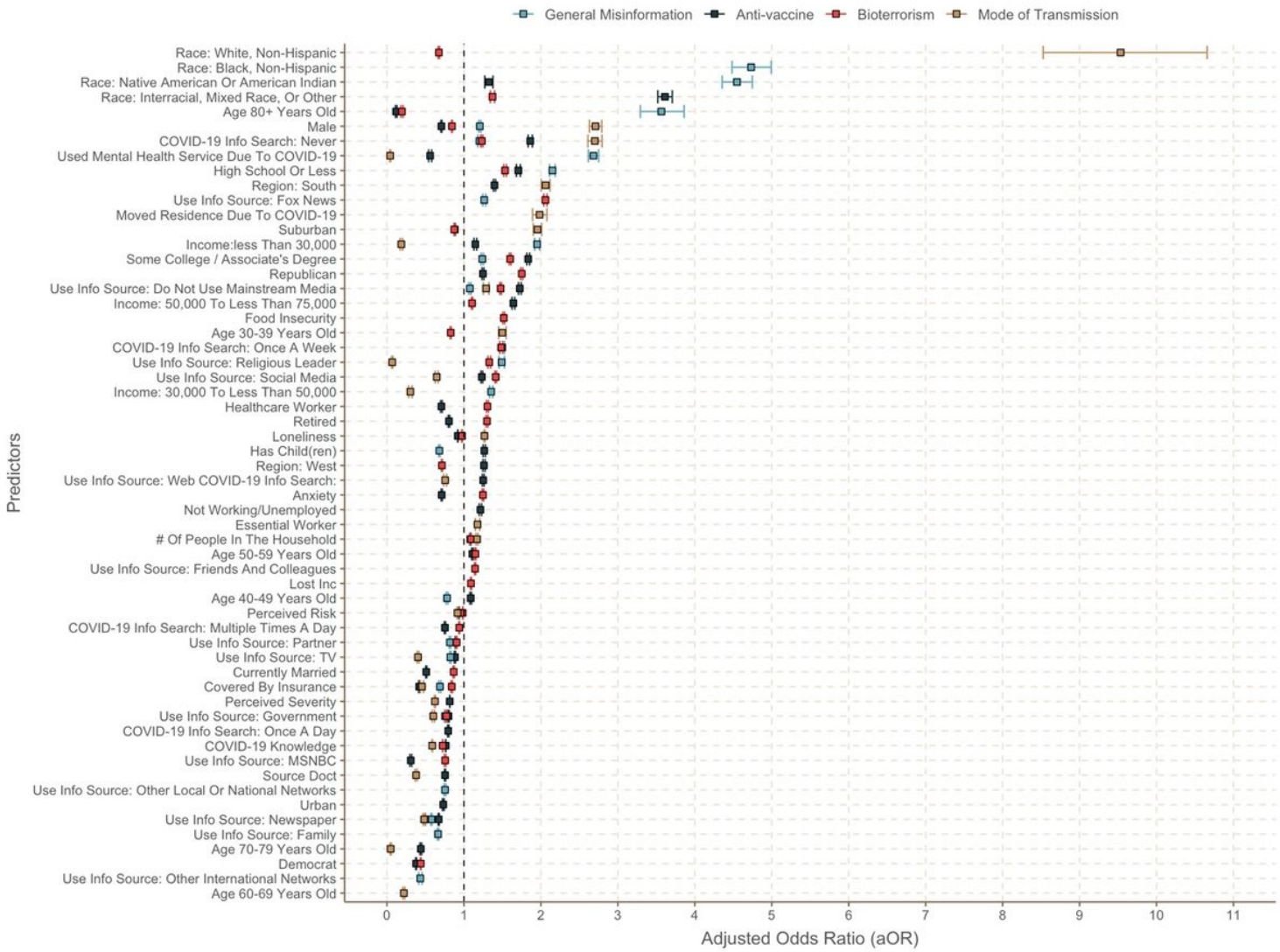

Figure 2

Factors associated with adherence to four different types of COVID-19 misinformation (aggregated plot)

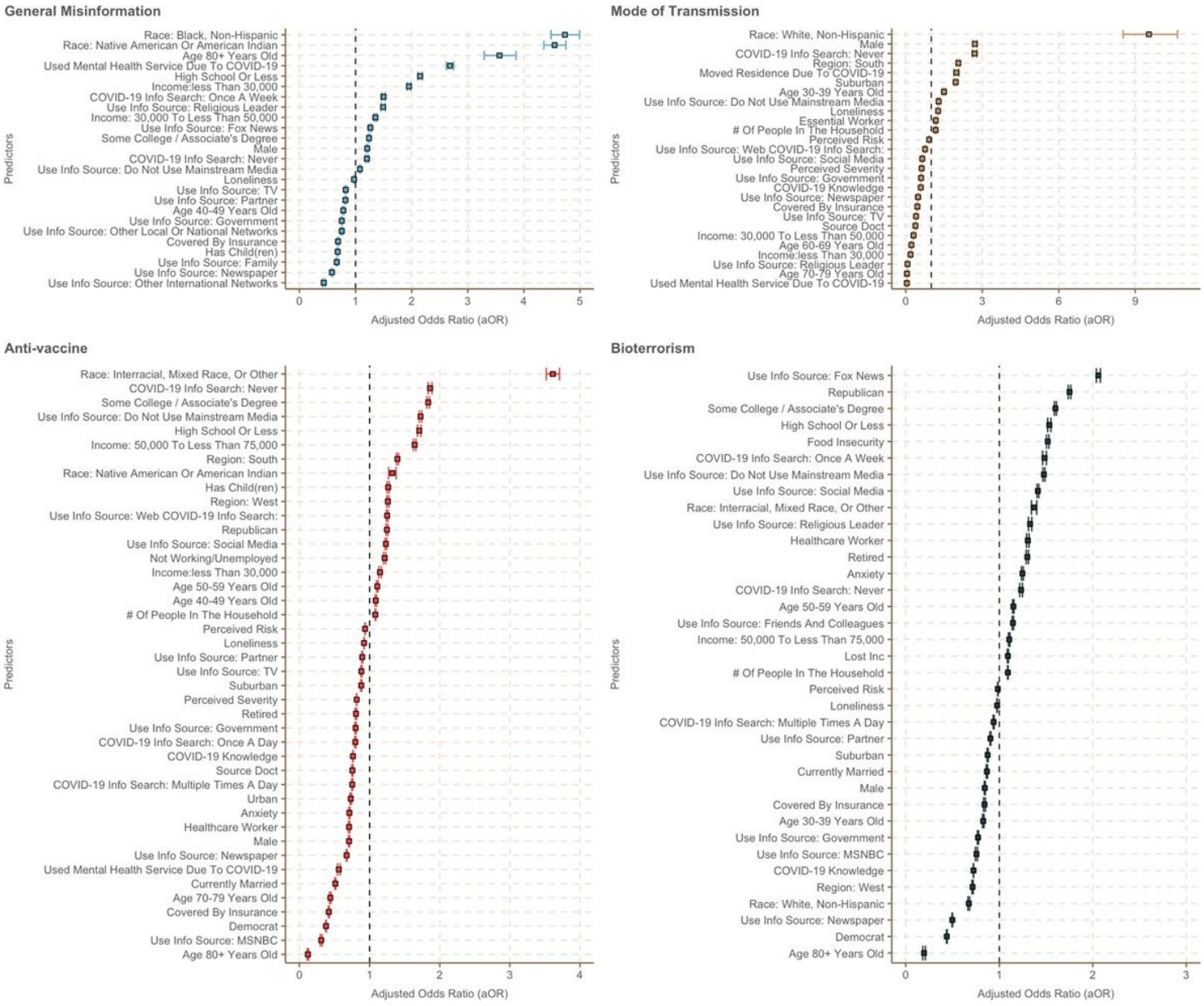

Figure 3

Factors associated with the adherence to four different types of COVID-19 misinformation (expanded plot)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.docx
- Additionalfile3.docx