



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original Research

Sociodemographic predictors of COVID-19 vaccine acceptance: a nationwide US-based survey study

Prithish Mondal ^{a, *}, Ankita Sinharoy ^b, Lilly Su ^c^a Division of Pediatric Pulmonology, Department of Pediatrics, Penn State College of Medicine, Hershey, PA, USA^b Penn State University, Harrisburg, PA, USA^c Penn State College of Medicine, Hershey, PA, USA

ARTICLE INFO

Article history:

Received 27 February 2021

Received in revised form

15 June 2021

Accepted 21 July 2021

Available online 29 July 2021

Keywords:

COVID-19

COVID-19 vaccine

Prediction model

Sociodemographic predictors

Machine learning

Vaccine hesitancy

ABSTRACT

Objectives: Acceptance of COVID-19 vaccination is attributable to sociodemographic factors and their complex interactions. Attitudes towards COVID-19 vaccines in the United States are changing frequently, especially since the launch of the vaccines and as the United States faces a third wave of the pandemic. Our primary objective was to determine the relative influence of sociodemographic predictors on COVID-19 vaccine acceptance. The secondary objectives were to understand the reasons behind vaccine refusal and compare COVID-19 vaccine acceptance with influenza vaccine uptake.

Study design: This was a nationwide US-based survey study.

Methods: A REDCap survey link was distributed using various online platforms. The primary study outcome was COVID-19 vaccine acceptance (yes/no). Sociodemographic factors, such as age, ethnicity, gender, education, family income, healthcare worker profession, residence regions, local healthcare facility and 'vaccine launch' period (pre vs post), were included as potential predictors. The differences in vaccine acceptance rates among sociodemographic subgroups were estimated by Chi-squared tests, whereas logistic regression and neural networks computed the prediction models and determined the predictors of relative significance.

Results: Among 2978 eligible respondents, 81.1% of participants were likely to receive the vaccine. All the predictors demonstrated significant associations with vaccine acceptance, except vaccine launch period. Regression analyses eliminated gender and vaccine launch period from the model, and the machine learning model reproduced the regression result. Both models precisely predicted individual vaccine acceptance and recognised education, ethnicity and age as the most important predictors. Fear of adverse effects and concern with efficacy were the principal reasons for vaccine refusal.

Conclusions: Sociodemographic predictors, such as education, ethnicity and age, significantly influenced COVID-19 vaccine acceptance, and concerns of side-effects and efficacy led to increased vaccine hesitancy.

© 2021 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

Introduction

The World Health Organisation (WHO) declared COVID-19 as a pandemic in March 2020,¹ and by June 2021, 173 million cases of COVID-19, including 3.74 million deaths, had been reported worldwide.² The preventive measures to contain COVID-19 were primarily unsuccessful in 2020, and the United States faced a

third wave of the pandemic in late 2020.³ Even countries such as India, which enjoyed initial success against the pandemic, have experienced sudden and overwhelming waves of cases and deaths as a result of a different coronavirus strains.⁴ Under these circumstances, mass immunisation is the principal strategy to manage the pandemic.⁵ Since the launch of COVID-19 vaccines, proimmunisation campaigns have been launched in the United States, and by the end of May 2021, most coastal states had achieved >50% adult immunisation. However, several states in the southeast region of the United States still face major vaccine hesitancy, with >60% of the adult population remaining unimmunised (as of June 2021).⁶

* Corresponding author. Division of Pediatric Pulmonology, Penn State College of Medicine, 500 University Drive, P.O. Box 850, Hershey, PA, 17033, USA. Tel.: +717 531-0003, x287234.

E-mail address: pmondal@pennstatehealth.psu.edu (P. Mondal).

Several sociodemographic factors could influence the acceptance of COVID-19 vaccines; among these factors, ethnic disparity has been well recognised.⁷ Although Asian individuals have the highest preference for COVID-19 vaccines, African American communities are usually more hesitant than other ethnicities.⁷ A higher education level and income, female gender and old age were also associated with a higher likelihood of vaccine acceptance.^{8,9} Attitudes towards COVID-19 vaccines also differ between US states. Malik et al. reported that residents of Rocky Mountain states (Denver region) had a 25% higher chance of vaccine acceptance than the Great Lake areas (Chicago region).¹⁰ Healthcare workers (HCWs) are also facing this dilemma, with a recent report showing that only 36% of HCWs were very confident about receiving the COVID-19 vaccine.¹¹ Vaccine hesitancy may follow trends, and this was identified as one of the top threats to global health by the WHO.¹² A mutually lower acceptance of influenza and COVID-19 vaccines has been reported in African American individuals compared with other ethnicities.¹⁰ In addition, a higher perceived risk of COVID-19 is a new concept and is recognised to increase the likelihood of COVID-19 vaccine acceptance.^{13–15}

At least 70% of the population needs to be vaccinated to build herd immunity against COVID-19.¹⁶ A recent survey reported that only 67% of the US population were willing to receive the COVID-19 vaccine.¹⁰ Peoples' perception of COVID-19 is evolving, as the situation with COVID-19 is frequently changing. Around mid-October 2020, a new wave of COVID-19 hit the United States, and the message was clear – COVID-19 is here to stay.^{17,18} In November 2020, Pfizer announced that their COVID-19 vaccine was 95% effective in disease prevention, and subsequently, the vaccine was launched.¹⁹ Soon after, social media platforms were flooded with reports of adverse effects.²⁰ Consequently, people became more sceptical about this COVID-19 vaccine.⁷ Amidst the surge of information and misinformation, it is crucial to acknowledge people's concerns about vaccination. An aggressive mass campaign targeting vaccine-hesitant populations, focusing on the risks versus the benefits of vaccination, is needed to gain trust and counter apprehension.^{11,13} However, the majority of population-based studies on COVID-19 vaccine acceptance were conducted before the launch of the vaccines.^{9,21} This study, which recruited a considerable number of study participants during the postlaunch period, is relevant and will help identify target populations for the policymakers and stakeholders.

We designed a survey study with the primary aim of building a prediction model to determine the relative influence of socio-demographic predictors on the decision of vaccine acceptance. Our secondary aims were to understand the reasons for COVID-19 vaccine refusal and to compare acceptance of COVID-19 vaccine with influenza vaccine uptake.

Methods

Study design, outcome measures, predictors and effect modifiers

A nationwide US-based cross-sectional survey study was conducted between May 2020 and January 2021. The survey questionnaire was created in REDCap (see [Supplemental file 1](#)). The study participants were grouped based on the following socio-demographic determinants (see [Table 1](#)): age (five groups in chronologically ascending order), gender, ethnicity (White, African American, Hispanic, Asian and others), level of education (five levels in ascending order) and family income (four groups defined in ascending order). Fifty US states were consolidated into nine divisions following the US census bureau's guidelines ([Table 1](#)).²² COVID-19 vaccine acceptance was the primary outcome, and socio-demographic factors were examined as potential predictors ([Table 1](#)). Study participants' professions (HCW vs non-HCW) and

satisfaction with available healthcare facilities (on a scale of 1–5) were also included in the survey questionnaire.

Pfizer announced success with their COVID-19 vaccine in November 2020, and the worst peak of the pandemic hit the United States around the same time.¹⁹ We chose 1 November 2020 as the time point to distinguish between pre- and post-vaccination launches. HCW profession, satisfaction with healthcare access and vaccine launch timeline were added to the study as effect modifiers (see [Table 2](#)).

Pennsylvania State University institutional review board approved the study protocol. Our primary recruitment tool was a web-based portal (Researchmatch²³), which includes 150,000 volunteers nationwide, representing all ages, ethnicities and genders, whom we could contact randomly. Researchmatch helped us to mitigate any selection bias. Additional recruitment platforms included Studyfinder, social media platforms, such as Facebook, and bulk email invitations to Penn State Health patients.²⁴

Data sets and variables

We used statistical software, SAS 9.4 and SPSS 27.^{25,26} The study population was primarily categorised into the following two groups: 'vaccine compliant' and 'vaccine hesitant'. We computed a new variable 'perceived risk' based on the following parameters: (a) dichotomous variables (yes vs no): age >60 years, HCW, family member diagnosed with COVID-19; and (b) stress (in 1–5 Likert scale) related to potential infection, the risk of severe disease, lack of hospital facility and COVID-19 prevalence in the US state of residence. Cronbach's alpha was used to test internal consistency among the variables contributing to the 'perceived risk'. The perceived risk scores were consolidated by 'factor reduction' into a nominal variable (range: 0–10). We also included questions to understand the reasons for refusal of the COVID-19 vaccine, such as concern with potential adverse effects or doubt about its efficacy and necessity. Finally, we enquired about the preferred source of information such as television, social media, Centers for Disease Control and Prevention (CDC), state health websites and personal communication.

Sample size estimation

Considering the US adult population of 255,200,373²⁷ and assuming that 50% of this population would be happy to receive a COVID-19 vaccine, at least 1068 survey participants were required to achieve 95% confidence level (margin of error: 3%).

Statistical analyses

We used Chi-squared tests to estimate the differences in vaccine acceptance among the subgroups of various predictors. Perceived risk scores were compared between vaccine-compliant groups and vaccine-hesitant groups using *t*-tests.

Logistic regression and neural network, a machine learning tool, were used to build prediction models to estimate the likelihood of COVID-19 vaccine acceptance based on socio-demographic determinants. Both methods listwise removed the missing data. We randomly selected 80% of the subjects to train (build) and the remaining 20% to test both of the models.

A stepwise backward logistic regression method was used. The regression model was initiated with nine predictors, which eliminated variables of least significance in a stepwise manner, until the optimum model was attained.

During the multilayer perception neural network, all predictors were incorporated as an input layer (factors). Subsequently, multiple hidden layers (nodes) were generated by default (see

Table 1
Distribution of sociodemographic predictors and the rate of COVID-19 vaccine acceptance.

Sociodemographic predictors (N ^a)		Vaccine acceptance	Chi-squared (P-value) ^b
Gender (N = 2934)	Female (N = 2215)	79.7%	$\chi^2 = 10.42 (P = 0.001)$
	Male (N = 719)	85.1%	
Family income (N = 2968)	Lower middle class (N = 432)	71.99%	$\chi^2 = 73.72 (P = 0.001)$
	Middle class (N = 1725)	79.25%	
	Upper middle class (N = 752)	90.43%	
	Upper class (N = 59)	89.83%	
US regions (N = 2914)	East North Central (N = 430)	81.6%	$\chi^2 = 28.48 (P = 0.001)$
	East South Central (N = 215)	81.9%	
	Mid-Atlantic (N = 716)	77.4%	
	Mountain (N = 201)	89.6%	
	New England (N = 107)	87.9%	
	Pacific (N = 336)	82.1%	
	South Atlantic (N = 520)	79.2%	
	West North Central (N = 190)	87.9%	
	West South Central (N = 199)	81.9%	
	Healthcare worker (N = 2966)	Yes (N = 542)	
No (N = 2424)		82.01%	
Ethnicity (N = 2874)	Asian (N = 129)	91.47%	$\chi^2 = 73.99 (P < 0.001)$
	African American (N = 197)	64.97%	
	Hispanic (N = 203)	71.92%	
	White (N = 2217)	72.66%	
	Others (N = 128)	84.12%	
Age groups (N = 2962)	I (18–24 years; N = 175)	84.00%	$\chi^2 = 59.21 (P < 0.001)$
	II (25–44 years; N = 950)	77.89%	
	III (45–60 years; N = 783)	75.35%	
	IV (61–70 years; N = 645)	86.05%	
	V (>70 years; N = 409)	90.71%	
Education (N = 2946)	High school (N = 354)	69.77%	$\chi^2 = 84.85 (P < 0.001)$
	Undergraduate (N = 823)	78.98%	
	Graduate (N = 624)	77.08%	
	Masters (N = 709)	88.86%	
	PhD/professional (N = 436)	88.99%	
Vaccine launch period (N = 2978)	Prelaunch (N = 519)	78.2%	$\chi^2 = 2.08 (P = 0.14)$
	Postlaunch (N = 2459)	80.5%	

The variance in COVID-19 vaccine acceptance within each group was estimated with analysis of variance and Chi-squared test (for dichotomous variables).

^a N represents the number of survey respondents who belong to that category.

^b P-value <0.05 was considered statistically significant.

[Supplemental file 2](#)). The hidden nodes contained functions of each input layer. The model architecture (e.g. hidden layers of numbers) was automatically selected by the tool. Finally, outcome layers (vaccine acceptance: yes vs no) were computed with each hidden layer's relative contribution. Neural network generated a list ranking the predictors in order of importance. The concordance between measured versus estimated values represented the strength of the prediction model.

Results

Sociodemographic distribution

We received 4183 responses, 449 of whom were from outside the United States. Of the remaining 3734 participants, we only

included those who completed all three sections of the survey questionnaire and answered the question that determined the primary outcome of vaccine acceptance. Finally, 2978 participants were included in the study.²⁸ The majority of participants were female (75.1%) and of White ethnicity (77.1%). The predominant age group was group II (25–44 years: 32.1%), followed by group III (45–60 years: 26.4%) and group IV (61–70 years: 21.8%; see [Table 1](#)). Survey respondents were relatively homogeneously distributed in terms of educational level, with >20% of participants belonging to each of the groups of undergraduates, graduates or masters. More than half of the study participants were from the 'middle class', followed by the 'upper' and 'lower middle class'. We received survey responses from all the US regions, ranging from 24.6% (Mid-Atlantic) to 3.7% (New England). Eight out of the nine regions had more than 6% share of the participant pool.

Table 2
Source of COVID-19-related knowledge among the vaccine-compliant and vaccine-hesitant participants and their differences were compared with the Chi-squared test.

Source of information	Vaccine compliant	Vaccine hesitant	Pearson Chi-squared (P-value) ^a
Television	67.1%	64.5%	1.31 (0.252)
Friends and family	39.7%	38.3%	0.368 (0.544)
CDC website	61.7%	54.3%	10.652 (0.001)
Facebook	23.6%	24.1%	0.072 (0.789)
City or state website	68.2%	56.0%	30.28 (<0.001)
Local doctor/hospital	32.1%	31.6%	0.053 (0.818)
Scientific magazine	23.7%	19.0%	5.891 (0.015)

Every survey respondent could select multiple options.

CDC, Centers for Disease Control and Prevention.

^a P-value <0.05 was considered statistically significant.

Predictors of COVID-19 vaccine acceptance

Vaccine acceptance rate among the study participants was 81.1%.

Age

The oldest participants (aged >70 years) had the highest vaccine acceptance rate (90.7%), followed by age group IV and I, whereas participants in age group III, followed by II, had the lowest vaccine acceptance rate (Table 1). A Chi-squared test demonstrated that the age group variance of vaccine acceptance was statistically significant ($P < 0.001$).

Ethnicity

Asian people (91.5%) had the highest vaccine acceptance rate, followed by White (84.1%), Hispanic (71.9%) and African American (65.0%) ethnic groups (Table 1). A Chi-squared test indicated significant differences in vaccine acceptance among different ethnic groups ($P < 0.001$).

Gender

Male participants reported a higher preference for COVID-19 immunisation compared with females (85.1% vs 79.7%, respectively) based on the Chi-squared test ($P = 0.001$).

Education level

Vaccine acceptance rate steadily increased with a higher level of education, except for a very similar acceptance rate between undergraduates (79.0%) and bachelors (77.1%; Table 1). A Chi-squared test demonstrated that vaccine acceptance among various education groups was significantly different ($P < 0.001$).

Family income

Participants from the lower middle class (group I) had the lowest vaccine acceptance rate (72.0%), followed by the middle class (79.3%). The upper middle and upper-class participants had an acceptance rate above 88% (Table 1). Vaccine acceptance significantly differed by family income group ($P < 0.001$).

Healthcare workers

HCWs had lower vaccine acceptance than non-HCWs (Table 1). There was a significant association between HCWs (yes vs no) and vaccine preference (yes vs no; $P = 0.02$).

US region

COVID-19 vaccine acceptance varied significantly between different regions of residence ($P = 0.001$), ranging from 89.6% (Mountain) to 77.4% (Mid-Atlantic; Table 1, Fig. 1).

Healthcare access

Vaccine-compliant participants were more concerned with available healthcare facilities than vaccine-hesitant groups (4.21 ± 0.92 vs 3.95 ± 1.08 ; t -tests: $P < 0.001$).

Vaccine launch

In total, 519 and 2459 subjects were recruited pre- and post-vaccine launch, respectively (Table 1). Vaccine acceptance rate did not differ between pre- and post-vaccine launch (78.2% vs 80.5%; $P = 0.14$).

Perceived risk

Variables contributing to the perceived risk demonstrated reliability (Cronbach's alpha = 0.632, which was acceptable since the number of variables was <10).²⁹ Vaccine-compliant individuals

reported higher perceived risk (5.17 ± 1.93) compared with vaccine-hesitant individuals (4.83 ± 2.36 ; t -test: $P = 0.001$).

Source of COVID-19-related information

Both vaccine-compliant and vaccine-hesitant groups had equivalent reliance on television, social interaction and social media (e.g. Facebook) to acquire COVID-19-related information (no statistical significance in Chi-squared test, Table 2). However, a significantly higher number of vaccine-compliant participants gained COVID-19 information from the CDC and other official health websites (Table 2).

Prediction models

The regression model (Table 3) was statistically significant ($\chi^2 = 1219.450$ [degree of freedom {df} = 29, $N = 2378$], $P < 0.001$), and all predictors significantly contributed to the model (Table 4, Fig. 2) except HCW, gender and vaccine launch period. The neural network method had an advantage over the regression model because it demonstrated the relative influence of individual predictors on vaccine acceptance. Education, ethnicity and age were the three most significant predictors for vaccine acceptance in both regression and neural network models (Table 4). Similar to the regression model, the neural network algorithm also recognised HCW, gender and vaccine launch period as the least important predictors of vaccine acceptance (Table 4).

Model performance and cross-validation

Logistic regression

The percentage of correct predictions in the training and testing models were 81.7% and 80.8%, respectively. Cox and Snell's pseudo R^2 estimate was 0.41, demonstrating a good model fit.

Neural network

The percentage of correct predictions in the training and testing models were 80.9% and 81.3%, respectively.

Concern about the vaccine

Potential adverse effects (84.2%), followed by doubt about vaccine efficacy (44.4%), were the primary reasons reported by the vaccine-hesitant group. Even 51.0% of the vaccine-compliant participants were concerned about adverse effects; however, only 5.4% doubted the efficacy of the vaccine. We found that middle-aged groups were more doubtful about vaccine efficacy, rather than adverse effects, compared with the oldest population group. In contrast, African American individuals were more concerned about adverse effects than vaccine efficacy compared with Asian and White ethnic groups. Females were also more worried about potential adverse effects than males.

Influenza vaccine

A significant association was found between influenza vaccine uptake and COVID-19 vaccine acceptance ($\chi^2 = 316.6$ [df = 4, $N = 2976$], $P < 0.001$), as 90.7% of vaccine-compliant participants had received the influenza vaccine in previous years, compared with only 61.0% of vaccine-hesitant participants. Similar to the COVID-19 vaccine, acceptance of the influenza vaccine was lowest among African American individuals (78.7%) compared with the White ethnic group (86.8%). We divided nine US regions into high, intermediate and low acceptance groups for both the COVID-19 and influenza vaccines separately. When compared, six of the nine

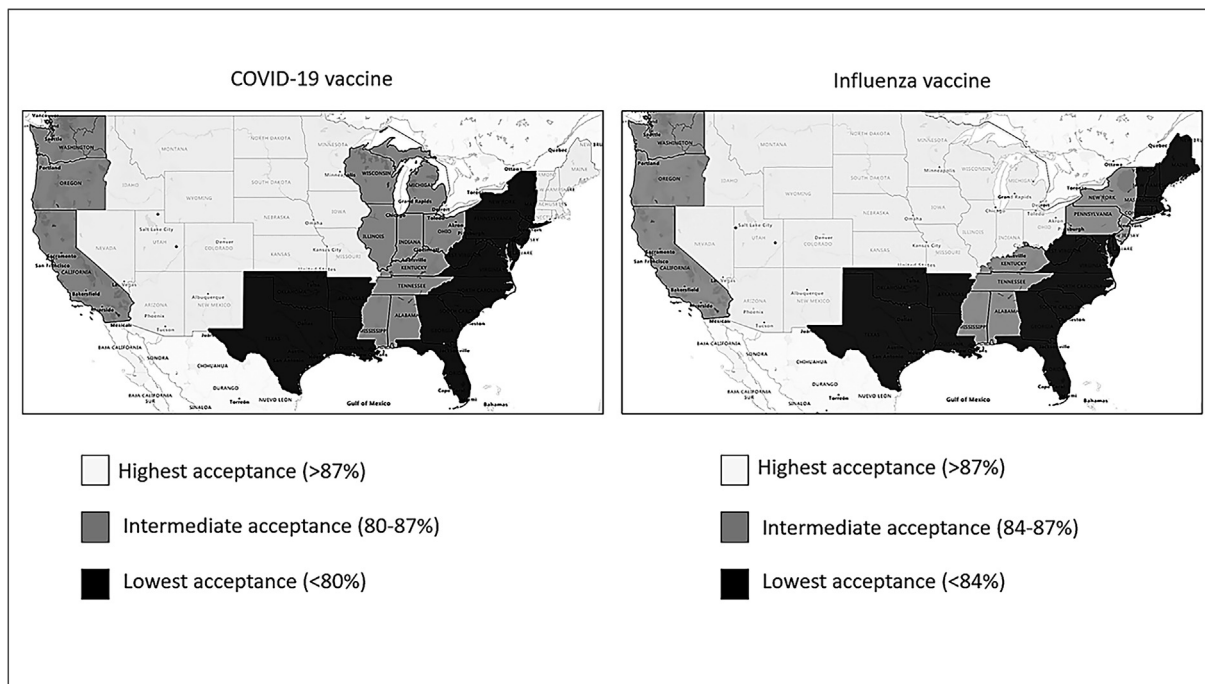


Fig. 1. Geographical heterogeneity of rate of acceptance of COVID-19 vaccine vs influenza vaccine. The US states were grouped into nine regions (census region). The regions were further ranked as high, intermediate and low categories separately based on the both the vaccines' acceptance rates. Six of nine regions fell into identical categories for both the vaccines.

regions were matched in terms of acceptance between the two vaccines (Fig. 1).

Discussion

This study demonstrates that sociodemographic factors, such as education, ethnicity and age, were the predominant predictors of acceptance of COVID-19 vaccine, and the models built on socio-demographic factors could precisely forecast COVID-19 vaccine acceptance. In total, COVID-19 vaccine acceptance rate was 81% in our study population, which is an encouraging sign because several earlier studies reported a lower acceptance rate (60–70%).¹⁰ The vaccine acceptance rate was unchanged among our study population even after the vaccine announcement, which is an interesting observation. Perhaps, several conflicting factors, such as the pro-vaccination campaign, reported adverse effects, the third wave of the pandemic and the arrival of newer strains, created doubt in peoples' minds. Nonetheless, the acceptance rate among our study participants reproduced the recent trend reported by Kreps et al.³⁰

Asian and African American ethnic groups had the highest difference (26.5%) in vaccine acceptance among the various ethnic groups, which is similar to a previous report.¹⁰ In addition to ethnicity, the neural network model selected age-wise distribution as a key predictor. Although the low acceptance rate among African American individuals has recently been highlighted, a sceptical middle-aged group should also be considered from a strategic perspective. Participants aged >70 years were highly motivated to receive the vaccine as they recognised COVID-19 as a threat to life, whereas some of the participants of younger age groups (II and III) perhaps considered COVID vaccination to be optional. This observation underscored the influence of perceived risk to encourage vaccine acceptance and explained why vaccine compliance was higher among older people and participants concerned with access to local hospitals. Higher education levels and family income also had a major influence on vaccine acceptance. Our female study

participants were more sceptical about the COVID-19 vaccine compared with male participants. However, recent reports suggest that women have been more proactive in receiving the vaccine, although they have also been reporting more adverse effects than men.^{31,32}

Vaccine acceptance rate also varied significantly among the US regions. Interestingly, the regional trend of COVID-19 vaccine acceptance somewhat followed that of the influenza vaccine uptake (Fig. 1). Respondents who refused the influenza vaccine were also hesitant to receive the COVID-19 vaccine, especially African American individuals, which has been previously reported.¹⁰ However, the degree of association between COVID-19 vaccine and influenza vaccine acceptance was more substantial among our study participants, irrespective of ethnicity, compared with the earlier report.¹⁰

As of 3 June 2021, 2.1 billion COVID-19 vaccinations have been administered worldwide.³³ Despite the claim of postvaccination deaths in social media, scientific investigations have not revealed any association between COVID-19 vaccination and these incidents.^{34,35} We believe that this message needs to reach all sections of the community, as potential adverse effects was a major concern, irrespective of vaccine acceptance. Messenger RNA vaccines have been reported to achieve 95% success in preventing serious COVID-19.³⁶ However, a considerable number of vaccine-hesitant people (44.4% in our study) remain sceptical about its efficacy primarily because of the speed of vaccine development and launch.³⁷ Ethical heterogeneity is an important observation from a strategic point of view, and alleviating fear among the African American ethnic group has been the predominant focus of pro-vaccination campaigns.³⁸ Vaccine information targeted towards the middle-aged population should also be a priority, addressing the efficacy concern (i.e. that with the technique, expertise and thousands of volunteers available, a speedy vaccine rollout was commendable, not concerning).³⁹

The question remains of how best to effectively target vaccine-hesitant populations. Interestingly, the vaccine-compliant and

Table 3
Odds ratio estimates of sociodemographic predictors by stepwise backward logistic regression.

Sociodemographic predictors	Odds ratio estimates			Estimate	Standard error	Wald Chi-squared ^a	P-value ^b
	Point estimate	95% Wald confidence limits					
Age (years)							
18–24 vs >70	1.542	0.817	2.911	0.433	0.324	1.784	0.18
25–44 vs >70	2.494	1.580	3.938	0.914	0.233	15.390	<0.001
45–60 vs >70	2.730	1.730	4.308	1.004	0.233	18.604	<0.001
61–70 vs >70	1.700	1.055	2.738	0.531	0.243	4.759	0.03
Education level							
Highschool vs PhD/professionals	3.011	1.908	4.751	1.102	0.233	22.412	<0.001
Undergraduate vs PhD/professionals	1.700	1.119	2.582	0.531	0.213	6.196	0.01
Graduate vs PhD/professionals	2.155	1.407	3.300	0.768	0.217	12.457	<0.001
Masters vs PhD/professionals	0.860	0.548	1.349	−0.151	0.230	0.432	0.51
Ethnicity							
Asian vs White	0.556	0.271	1.139	−0.588	0.366	2.576	0.11
Black vs White	2.043	1.373	3.041	0.714	0.203	12.407	<0.001
Hispanic vs White	1.521	1.013	2.282	0.419	0.207	4.092	0.04
Others vs White	1.767	1.095	2.853	0.569	0.244	5.427	0.02
Family income							
Lower middle class vs upper class	1.646	0.612	4.432	0.499	0.505	0.974	0.32
Middle class vs upper class	1.627	0.624	4.241	0.487	0.489	0.992	0.32
Upper middle class vs upper class	0.797	0.296	2.142	−0.227	0.505	0.203	0.65
Healthcare worker	0.718	0.541	0.953	−0.331	0.144	5.279	0.02
US regions							
East North Central vs West South Central	0.642	0.392	1.053	−0.442	0.252	3.077	0.08
East South Central vs West South Central	0.675	0.384	1.189	−0.393	0.288	1.853	0.17
Mid-Atlantic vs West South Central	0.785	0.496	1.240	−0.243	0.234	1.080	0.30
Mountain vs West South Central	0.363	0.189	0.699	−1.013	0.334	9.202	0.002
New England vs West South Central	0.272	0.110	0.669	−1.303	0.459	8.042	0.005
Pacific vs West South Central	0.791	0.473	1.323	−0.235	0.262	0.799	0.37
South Atlantic vs West South Central	0.820	0.512	1.315	−0.198	0.241	0.676	0.41
West North Central vs West South Central	0.474	0.248	0.904	−0.748	0.330	5.130	0.02
Hospital satisfaction							
Strongly disagree vs strongly agree	1.872	0.950	3.688	0.627	0.346	3.282	0.07
Disagree vs strongly agree	1.804	1.138	2.860	0.590	0.235	6.294	0.01
Neutral vs strongly agree	1.486	1.033	2.137	0.396	0.186	4.549	0.03
Agree vs strongly agree	1.052	0.813	1.361	0.050	0.132	0.147	0.70

^a Relative contribution of a predictor to the model.

^b P-value <0.05 was considered statistically significant.

vaccine-hesitant groups had similar preferences for obtaining vaccine-related information from the television, community and social media, such as Facebook. However, the vaccine-hesitant group had significantly less inclination towards institutions such as CDC or state health authorities. Thus, to reach these segments of the population, the approach needs to be customised at a personal level, and perhaps role-modelling may help.^{40,41} The similarity in influenza and COVID-19 vaccine acceptance trends indicates that people who refused influenza vaccine were also likely to be COVID-19 vaccine hesitant. Thus, strategies to promote influenza vaccine

uptake could also be effective in increasing COVID-19 vaccine acceptance.⁴² In addition, we analysed factors influencing vaccine hesitancy, which should be considered in future informed strategies.

This study has a few limitations that need to be acknowledged. Questions regarding vaccine-related concerns were added to the survey halfway through the recruitment period; thus, only half (n = 1328) of the participants answered these questions. As this was an internet-based survey, a section of the computer-averse or non-native English-speaking population were excluded. Also, the results from this study may have been influenced by selection bias;

Table 4
Significance of each of the nine predictors to the logistic regression model (stepwise backward), based on Wald Chi-squared and P-values.

Predictors	Logistic regression			Neural network	
	df	Wald Chi-squared	P-value ^a	Relative importance	Rank
Education level	4	49.61	<0.001	16.60%	2
Ethnicity	4	31.87	<0.001	15.80%	3
Age	4	27.36	<0.001	18.80%	1
US region	8	22.51	0.004	15.20%	4
Family income	3	21.16	<0.001	10.30%	5
Healthcare access	4	11.60	0.02	9.90%	6
Healthcare worker	1	5.32	0.07	6.90%	7
Gender ^b	1	1.03	0.23	5.20%	8
Vaccine launch period ^b	1	2.93	0.60	1.30%	9

The stepwise backward regression excluded gender and vaccine launch period from the final regression model. The neural networks model ranked the predictors based on their relative importance.

^a P-value <0.05 was considered statistically significant.

^b Gender and vaccine launch period were excluded from the final regression model.

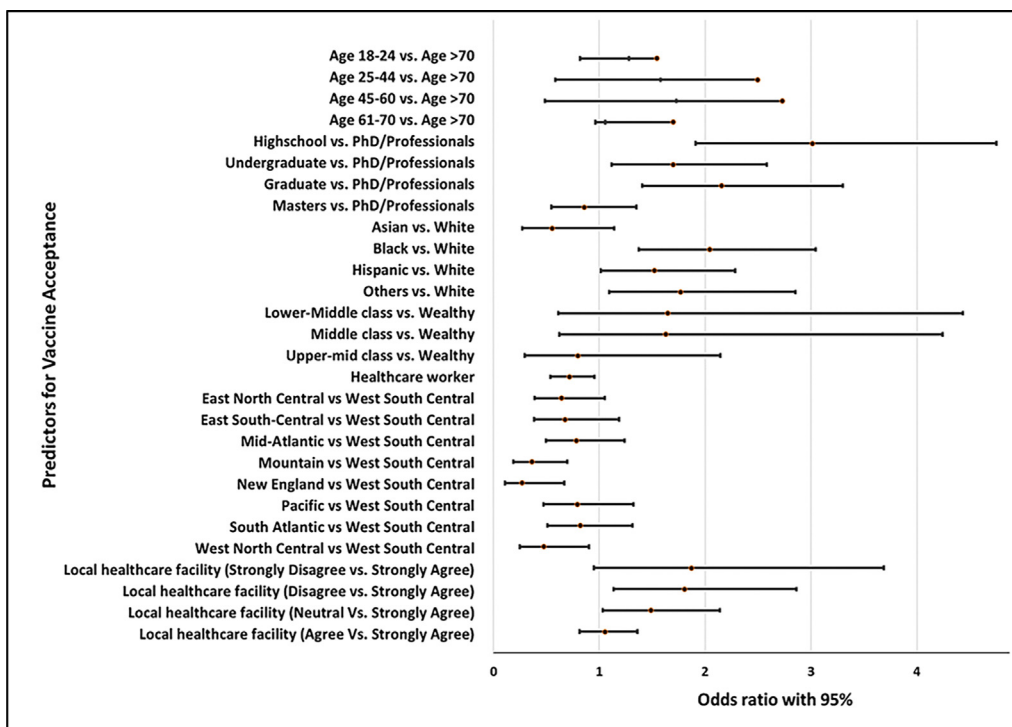


Fig. 2. Forest plot demonstrating the odds ratios (point estimate with 95% confidence interval) of vaccine acceptance among individual subgroups compared with the reference subgroup, within each predictor.

however, the primary recruitment portal (Researchmatch) was designed to mitigate that challenge and recruited participants from across the country.

Nonetheless, this study has several strengths. Only a few previous studies have reported a similar prediction model of COVID-19 vaccine acceptance based on relative importance and interaction of sociodemographic factors in the United States.¹⁰ Hopefully, this report will help to prioritise future strategies. Moreover, many of the study participants were recruited after the announcement of the COVID-19 vaccine, thus capturing a very recent trend (up to January 2021). Logistic regression and neural network models were consistent with prediction accuracy and top predictors, and both methods identified HCW, gender and vaccine announcement period as factors of limited significance. Finally, reproducing prior study results and successful cross-validation of both the models added reliability to this study.

To conclude, prediction models built on sociodemographic factors using logistic regression and neural network could accurately forecast the likelihood of COVID-19 vaccine acceptance. Both models recognised education, ethnicity and age as predictors of high importance and excluded HCW, gender and the vaccine launch period.

Author statements

Acknowledgments

The authors would like to acknowledge Ms Deaven Hough for helping the project with the promotion of the survey and participant recruitment.

Ethical approval

The institutional review board at Penn State College of Medicine approved the study protocol.

Funding

None declared.

Competing interests

The authors have no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.puhe.2021.07.028>.

References

- <https://www.who.int/news/item/29-06-2020-covidtimeline>.
- https://covid19.who.int/?gclid=Cj0KCQiArvX_BRCyARIsAKsnTxP8vFO0-p1vB-DCjvhDjO-M3HjD9S5M37XWtjezJGps3uthX6Glu58aAg9-EALw_wcB.
- <https://www.cidrap.umn.edu/news-perspective/2020/12/us-sees-record-number-covid-19-deaths-december>.
- <https://www.mohfw.gov.in/>.
- <https://theconversation.com/can-we-actually-learn-to-live-with-coronavirus-not-until-we-have-a-vaccine-147792>.
- <https://www.mayoclinic.org/coronavirus-covid-19/vaccine-tracker>. Accessed 06/02/2021.
- Latkin CA, Dayton L, Yi G, Konstantopoulos A, Boodram B. Trust in a COVID-19 vaccine in the US: a social-ecological perspective. *Social Science & Medicine* 2021;**270**:113684. 1982.
- <https://www.pewresearch.org/science/2020/12/03/intent-to-get-a-covid-19-vaccine-rises-to-60-as-confidence-in-research-and-development-process-increases/>.
- Lazarus JV, Ratzan SC, Palayew A, et al. A global survey of potential acceptance of a COVID-19 vaccine. *Nat Med* 2021;**27**(2):225–8.
- Malik AA, McFadden SM, Elharake J, Omer SB. Determinants of COVID-19 vaccine acceptance in the US. *EClinicalMedicine*. 2020;**26**:100495.
- Shekhar R, Sheikh AB, Upadhyay S, Singh M, Kottewar S, Mir H, Barrett E, Pal S. COVID-19 vaccine acceptance among health care workers in the United States. *Vaccines* 2021 Feb;**9**(2):119.
- <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>.

13. Caserotti M, Girardi P, Rubaltelli E, Tasso A, Lotto L, Gavaruzzi T. Associations of COVID-19 risk perception with vaccine hesitancy over time for Italian residents. *Soc Sci Med* 2021;113688.
14. Khubchandani J, Sharma S, Price JH, Wiblehauser MJ, Sharma M, Webb FJ. COVID-19 vaccination hesitancy in the United States: a rapid national assessment. *J Community Health* 2021;1–8.
15. McFadden SM, Malik AA, Aguolu OG, Willebrand KS, Omer SB. Perceptions of the adult US population regarding the novel coronavirus outbreak. *PLoS One* 2020;15(4):e0231808.
16. <https://health.clevelandclinic.org/how-much-of-the-population-will-need-to-be-vaccinated-until-the-pandemic-is-over/>. Accessed 06/03/2021.
17. <https://time.com/5903673/record-daily-coronavirus-cases/>.
18. <https://www.bloomberg.com/news/features/2021-01-28/covid-is-here-to-stay-pfizer-ceo-albert-bourla>.
19. <https://www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-announce-vaccine-candidate-against>.
20. <https://www.cnn.com/2020/12/01/tech/covid-19-vaccine-misinformation-social-media/index.html>.
21. Shekhar R, Sheikh AB, Upadhyay S, et al. COVID-19 vaccine acceptance among health care workers in the United States. *Vaccines*. 2021;9(2):119.
22. <https://www2.census.gov/geo/pdfs/reference/GARM/Ch6GARM.pdf>.
23. <https://www.researchmatch.org/>.
24. <https://studyfinder.psu.edu/>.
25. <https://support.sas.com/software/94/>.
26. <https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-27010>.
27. <https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html>.
28. Mondal P, Sinharoy A. The factors determining the preference for COVID-19 Vaccine. In: *Mendeley data*. Elsevier; 2021.
29. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ* 2018;48(6):1273–96.
30. Kreps S, Prasad S, Brownstein JS, et al. Factors associated with US adults' likelihood of accepting COVID-19 vaccination. *JAMA network open*. 2020;3(10):e2025594–e2025594.
31. <https://khn.org/news/article/gender-vaccine-gap-more-women-than-men-vaccinated-against-covid/>.
32. <https://www.nytimes.com/2021/03/08/health/vaccine-side-effects-women-men.html>.
33. https://ourworldindata.org/covid-vaccinations?country=OWID_WRL.
34. <https://www.dw.com/en/fact-check-are-covid-19-vaccines-causing-deaths/a-56458746>.
35. <https://www.cbsnews.com/news/norway-covid-19-vaccine-elderly-deaths-no-link/>.
36. <https://www.statnews.com/2021/02/02/comparing-the-covid-19-vaccines-developed-by-pfizer-moderna-and-johnson-johnson/>.
37. <https://www.pharmaceutical-technology.com/features/how-to-combat-vaccine-hesitancy-in-the-age-of-covid-19/>.
38. Akarsu B, Canbay Özdemir D, Ayhan Baser D, Aksoy H, Fidancı İ, Can-kurtaran M. While studies on COVID-19 vaccine is ongoing, the public's thoughts and attitudes to the future COVID-19 vaccine. *Int J Clin Pract* 2020:e13891.
39. <https://nortonhealthcare.com/news/reaching-those-skeptical-about-covid-19-vaccines/>.
40. <https://www.usatoday.com/story/news/nation/2020/12/24/covid-vaccines-here-how-persuade-people-get-them/3939114001/>.
41. Laine C, Cotton D, Moyer DV. COVID-19 vaccine: promoting vaccine acceptance. *Ann Int Med* 2021;174(2):252–3.
42. Bhat-Schelbert K, Lin CJ, Matambanadzo A, Hannibal K, Nowalk MP, Zimmerman RK. Barriers to and facilitators of child influenza vaccine—Perspectives from parents, teens, marketing and healthcare professionals. *Vaccine* 2012;30(14):2448–52.