Predicting COVID-19 Vaccination Intention: Using A Machine Learning Approach

Xin Zhang[1], Zhixin Kang[2], Guanlin Gao[3], Xinyan Shi[4]

Abstract

This study examines the factors that influence people's decision to be vaccinated and their intentions to vaccinate their children. We conducted a cross-sectional online survey that was distributed through the Lumbee Tribe of North Carolina from the end of July 2021 to the beginning of August 2021. The survey questions include sociodemographic questions, questions about people's testing behaviors, vaccination status and vaccination intentions, COVID-19 knowledge, attitude, and beliefs. We employed four different Machine Learning models and find that the Random Forest algorithm best predicts the self-vaccination intention with a 96% of accuracy and children vaccination intention with a 86.5% of accuracy, respectively. The results also show that testing frequency, social responsibility and paid leave are strong predictors of people's vaccination decisions for themselves and their children. Policymakers may consider shifting resources to benefit communities and the populations facing frequent testing needs as well as emphasizing the social responsibility of vaccination to amplify interventions in promoting vaccination intake.

Keywords: COVID-19; vaccination intention; testing; social responsibility.

Funding source: This project was funded by National Institute on Minority Health and Health Disparity. The award number is 3R01MD012767-04S1.

[1] Department of Mathematics and Computer Science, The University of North Carolina at Pembroke. Email: xin.zhang@uncp.edu
[2] Department of Economics and Decision Science, The University of North Carolina at Pembroke. Email: Zhixin.kang@uncp.edu
[3] School of Business, Chaminade University of Honolulu. Email: guanlin.gao@chaminade.edu
[4] Department of Economics and Decision Science, The University of North Carolina at Pembroke. Email: yan.shi@uncp.edu. Corresponding Author.

1. Introduction

Mass vaccination is essential for ending the COVID-19 pandemic and returning life to normal. Previous studies show that race (Bagasra, Doan, & Allen, 2021; Fridman, Gershon, & Gneezy, 2021; Mackey et al., 2021; Viswanath et al., 2021), age (Banik et al., 2021; Detoc et al., 2020; Sherman et al., 2021), gender (Ali et al., 2020; Latkin et al., 2021; Lin et al., 2021; Robinson, Jones, & Daly, 2020; Zintel et al., 2022), religion (Chu, Pink, & Willer, 2021; Olagoke, Olagoke, & Hughes, 2021), political views (Barrios & Hochberg, 2020; Brinson, 2022; Pink et al., 2021), income (Paul, Steptoe, & Fancourt, 2021; Solís Arce et al., 2021; Thunström et al., 2021), and medical insurance status (Barry, 2021; Ku, 2021) are significant determinants of vaccination intention. And concerns over vaccination safety (Kerr et al., 2021; Lin, Tu, & Beitsch, 2021; Petersen et al., 2020) and side effects (Kerr et al., 2021; Lewis, 2020; Wong et al., 2021) are major contributors to vaccination hesitancy. Furthermore, studies also show that people's knowledge level, attitude and belief about COVID-19, and their confidence about their knowledge level affect their preventative behaviors (Bailey, 2020; Gao & Li, 2022; Lee, Kang & You, 2021; Rahman et al., 2022). Furthermore, recent research found that people more likely vaccine themselves than their children, and people's beliefs and attitudes towards the COVID-19 vaccine are the main determinants of vaccination intention (Fernandes et al. 2021).

Despite ample studies on the connection between various factors and COVID-19 vaccination intentions (Al-Amer et al., 2022; Carrieri et al., 2021; Fernandes et al., 2021; Sherman et al., 2021; Huynh et. al., 2021; Wang et. al., 2022), no study has investigated how testing behavior may reflect one's tendency to adopt a preventive behavior such as vaccination. Research on testing has mainly focused on evaluating testing patterns and understanding the testing rate among different populations and different locations (Brandt et al., 2021; Goldstein and Burstyn, 2020; Levesque et al., 2013). Although this information is crucial in order to evaluate the existing testing infrastructure and identify the strength and weakness to help develop an effective public health response, little is known on how testing patterns and testing behaviors are related to vaccination pattern and vaccination behavior, especially in marginalized communities. This paper attempts to fill in this gap by examining testing related factors such as the frequency of getting tested, difficulties experienced to get tested, knowledge on testing location and how these factors relate to vaccination decision. Additional factors examined include sociodemographic characteristics such as age, gender, race, income, education, and employment status, insurance status, people's beliefs and knowledge about COVID-19, and people's confidence about their knowledge level. We study the predictive power of these factors to explain individuals' vaccination intention for both themselves and their children.

We propose a machine learning approach to predict the participants' and their children's vaccination status. The machine learning algorithms have been applied in the health context in screening, diagnosis, and predicting the probability of an event (Panch, Szolovits, and Atun, 2018; Zhang et. al. 2011). One of the recent relevant applications of machine learning algorithms in healthcare is in the development of models to analyze COVID-19 seriousness, mortality, recovery, testing and vaccination hesitancy/intention (Carrieri, Lagravinese, and Resce, 2021; Fernandes et al., 2021; Iwendi, 2020; Mellado et al., 2021; Muhammad et al. 2020; Ong et al., 2020; Ritonga et al., 2021). For instance, Fernandes et al. (2021) used an artificial neural network (ANN) machine learning model to predict individuals' COVID-19 vaccine intention and

found that the COVID-19 vaccine beliefs and attitudes play major role in determining one's vaccination intention. Carrieri, Lagravinese, and Resce (2021) used four machine learning algorithms to study vaccination hesitancy and found that the proportion of waste recycling and the employment rate to be the most powerful predictors of high vaccination hesitancy.

In this paper, we use four machine learning models to predict the survey respondents' and their children's vaccination status based on the respondents' testing history, sociodemographic characteristics, COVID-19 belief and attitude, and other factors. These models are Logistic regression (LR), Decision Trees (DT), Artificial Neural Networks (ANN), and Random Forests (RF). We find that the RF algorithm performs the best to predict both the self-vaccination and children vaccination intention. The RF model for self-vaccination predicts with an accuracy rate of 96%, a precision of 96%, and a true positive rate of 99.8%, while the RF model for children vaccination predicts with an accuracy rate of 86.5%, a precision of 86.2%, and a true positive rate of 99.3%, respectively.

The results of this study also show that testing frequency, social responsibility and paid leave are strong predictors for both self and children vaccination intention. Adults who are tested more frequently, believe vaccination is a social responsibility and have paid leave for vaccination from their employers are more likely to be fully vaccinated and vaccinate their children. In addition, having received a flu shot in the last 12 months and believing that 'I am or will be more relaxed once I can get the vaccine' have strong influence on self-vaccination decision and are revealed by both LR and ANN models.

Compared to previous studies using hypothetical COVID-19 vaccination intentions (Al-Amer et al., 2022; Huynh et. al., 2021; Sherman et al., 2021; Wang et. Al., 2022), this study examines additional factors that are important to predict vaccination behavior and fills in the gap in the literature by providing empirical evidence on how testing and certain vaccination attitudes/beliefs relate to one's vaccination status.

2. Survey Design and Methodology

We constructed a survey based on the National Institutes of Health (NIH) RADx-UP Common Data Elements (CDEs)[5]. The detailed information is provided in the appendix. The survey questionnaires fell into the following blocks

Socio-demographic questions: Participants were asked about their gender, age, race, highest education qualification, family composition, employment status, income, and area of residence.

COVID-19 testing questions: Participants were asked about the number of testings received, the reasons to get tested, testing locations, barriers to get tested, the difficulty of getting tested, and whether they were tested positive.

[5] RADx-UP is a part of the NIH Rapid Acceleration of Diagnostics (RADx®) initiative to help speed innovation in the development and implementation of COVID-19 testing. The NIH RADx-UP CDEs provide a standard set of study questions for COVID-19 testing studies.

COVID-19 vaccine knowledge questions: The knowledge of COVID-19 was measured by asking participants the following five questions: 1) "*The only people who can get severely ill from COVID-19 are the elderly and people with certain medical conditions*."; 2) "*People who have COVID-19 and no symptoms can still spread it to others*."; 3) "*If someone tests negative for COVID-19, this person can still get infected in the future*."; 4) "*COVID-19 vaccines help keep people from getting seriously ill even if they do get infected with COVID-19*."; 5) "*Fully vaccinated people do not need to be tested after being exposed to someone with COVID-19*."

Moreover, participants were asked about their confidence of how they answer the knowledge questions and how an average person answers these questions correctly. They were asked to rate their confidences on a 5-point scale ranging from "Not confident at all" to "Extremely confident."

COVID-19 information questions. Participants were asked about the sources of information to learn about COVID-19 vaccine, and how easy or difficult to obtain information about COVID-19 vaccines, to sign up for a COVID-19 vaccine, and to find a time and location.

COVID-19 vaccine behavior questions: Participants were asked about the factors that influence their vaccination decision, their vaccination status and plan for themselves, the reasons for the decision of no vaccination, and whether they received a flu shot in the last 12 months. The questions regarding participants' intention to vaccine their children include their plan to get their children vaccinated and whether their children received flu shots in the last 12 months.

COVID-19 vaccine attitude and belief: Participants were also presented with a series of four statements concerning their attitudes and beliefs about COVID-19. They were asked to rate these questions on a 5-point scale ranging from "strongly disagree" to "strongly agree." The statements include: "*I am or will be more relaxed once I can get the vaccine*"; "*I may not wear a mask after fully vaccinated*"; "*Once the COVID vaccine is given to most of the population, our lives will finally return to normal again*", and "*It is my responsibility to get vaccinated so that our whole community is protected.*"

The questionnaire was developed online using Qualtrics and was distributed through the Lumbee Tribe of North Carolina's official website from the end of July to the beginning of August, 2021. The survey respondents are either members of LTNC or friends and relatives of LTNC members. About 30% of respondents are from North Carolina and the rest of them are located in other parts of the country. This study is approved by the Institutional Review Board of the University of North Carolina, Pembroke (UNCP IRB).

3. Data Analysis

To explain the intention/behavior to receive a COVID-19 vaccine, we built two linear regression models: one to predict the participants' intention to vaccinate themselves, and the other to predict the participants' intention to vaccinate their children. Additionally, we used four machine learning tools to build prediction models for vaccination intention for adults as wells as for their children. We performed statistical analysis using Weka 3.8.6 for all the machine learning models. We follow the approach adopted in the literature of health decision making using

machine learning (ML) models (see Brnabic & Hess, 2021, for a review), and randomly divide the dataset as 90% (890 for adults and 619 for children) for training and 10% (99 for adults 69 for children) for out-of-sample testing set.

4. Results

Summary Statistics

Our sample consisted of 1003 participants of the general adult population, ranging from 18 to 72 years old (M = 32.27; SD = 8.08). The sample was comprised 56.13% of males and 41.97% of females, and 1.89% of other gender types. A total of 56.33% of participants are White, 11.76% are African Americans, 16.35% are Native Americans, and 15.55% reported to have more than one race or other. There are 36.19% of participants with college degree, 34.6% with some college or trade school education, 15.75% with graduate or professional degree, 11.47% with high school diploma or GED, and 1.99% of participants dropped out of high school.

Table 1 reports the summary statistics of the key variables. The *Fully vaccinated* includes those who have received all recommended doses of a COVID-19 vaccine, those who are partially vaccinated, and those who have a definitive plan of getting fully vaccinated. The *Child vaccination* variable takes a value of 1 if the respondents report that they have or plan to have their children vaccinated and 0 otherwise. It is interesting to see that there is a higher percentage of people that were willing to have their children vaccinated compared with their self-vaccination percentage, which is different from the findings by Fernandes (2021). *Testing* is the number of COVID-19 testings the respondent received. An average person has been tested for three times. All other variables are binary and self-explanatory. We oversampled indigenous populations (mean=0.1956, i.e., 19.65% of the total sample), including Alaska Natives, Native Americans, Native Hawaiians and other Pacific Islanders, with the majority being Native Americans. We also asked participants if they have experienced any difficulties in testing and vaccination, and if they believe it is a social responsibility to get vaccinated. Table 1 also reported the answers to these questions as percentages.

Table 1 Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Fully vaccinated	929	.757	.429	0	1
Child vaccination	695	.813	.390	0	1
Testing frequency	997	3.056	2.271	0	10
Asian	1208	.024	.153	0	1
Black	1208	.104	.305	0	1
Hispanic	1208	.035	.183	0	1
Indigenous	1208	.197	.397	0	1
White	1208	.499	.500	0	1
Age≥30	1208	.589	.492	0	1
Female	1208	.349	.477	0	1
Annual household income < $25,000	981	.201	.401	0	1
College degree or above	1002	.520	.500	0	1

Employed (full and part-time)	1208	.702	.458	0	1
Have insurance	963	.872	.334	0	1
Testing difficulty: If you wanted to get a COVID test, how difficult would it be? (N=1,000)					
Extremely difficult					8.00%
Difficult					31.00%
Not difficult					61.00%
Has it been difficult for you to get information about COVID vaccines? (N=992)					
Yes					77.72%
No					22.28%
Has it been difficult for you to sign up for a COVID vaccine? (N=963)					
Yes					34.58%
No					65.42%
Has it been difficult for you to find a convenient time and place to be vaccinated? (N=974)					
Yes					37.99%
No					62.01%
It is my responsibility to get vaccinated so that our whole community is protected. (N=1,000)					
Agree					65.10%
Disagree					34.90%

Multiple Regression Model Results

Table 2 presents the nested multivariate linear regression results for Self and Children Vaccination Intention. These models only include the predictors that showed a significant effect in the original models that include all the variables. The predictors that significantly predicted both self-vaccination and children vaccination intention consist of *testing frequency*, *indigenous*, *vaccine attitude (social responsibility)*, *vaccine reason: paid leave*, *flu shot*. Additional significant predictors for self-vaccination intention include *black*, I*ndegious*, *nowledge_correct* (the number of knowledge questions answered correctly), r*equired* (vaccination required by the employer), *vaccine attitude* (more relaxed once being vaccinated), *vaccine attitude* (once vaccine is given to most of the population, our life will return to normal), *vacc_sign_up* (the difficulty to sign up for a vaccine), *vacc_time_place* (the difficulty to find a convenient time and place to be vaccinated). For the model with the intention to vaccinate the children as the outcome variable, there are additional significant predictors including *insurance*, *information resource*: *medical_provider*, *information resource*: *faith leader*, and *vaccine_info* (the difficulty to get information about vaccine).

Table 2: Multiple Linear Regressions

	OR	**OR**

Variables	Self-vaccination	Children vaccination
test frequency	1.140*	1.235***
	(0.075)	0.073
vaccine attitude: social responsibility	7.830***	1.623***
	(3.514)	0.225
paid leave	7.122***	2.182**
	(3.408)	0.685
flu shot	0.236***	0.333***
	(0.070)	0.083
black	0.159***	
	(0.064)	
Indegious	0.119***	
	(0.037)	
knowledge_correct	0.641***	
	(0.087)	
required	1.946*	
	(0.670)	
vaccine attitude: more relaxed with vaccine	0.565**	
	(0.140)	
vaccine attitude: returns to normal with vaccine)	0.613**	
	(0.153)	
vacc_sign_up	0.523***	
	(0.116)	
vacc_time_place	1.549**	
	(0.335)	
vacc_info		0.671***
		(0.096)
insurance		4.341***
		(1.500)
medical provider		0.610*
		(0.156)
faith leader		0.214***
		(0.109)
constant	23.916***	2.283
	(21.032)	(1.715)
Observations	879	623
Pseudo R2	0.377	0.231

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Machine Learning Model Results

Table 3 provides the standard classification performance metrics for the four different machine learning algorithms to predict the vaccination tendency for adults. Overall, the RF and ANN models perform better than LR and DT models do. In particular, the RF model performs the best and overperforms the three other models measured by the following metrics: Accuracy (96.0%), Recall (99.8%), MCC (63.4%), True positive rate (99.8%), False positive rate (55.1%), Area under the ROC curve (95.9%), Area under the PRC curve (99.6%). This forecasting evaluation process contains 64 attributes and 989 records.

Table 3 Model's Performance to Predict Adult Vaccination Decision

	LR	DT	RF	ANN
Accuracy	91.2%	94.0%	96.0%	95.3%
Precision	95.8%	96.3%	96.0%	97.2%
Recall	94.7%	97.3%	99.8%	97.7%
F-Measure	95.2%	96.8%	63.4%	97.5%
MCC	37.0%	51.2%	63.4%	62.2%
True positive rate	94.7%	97.3%	99.8%	97.7%
False positive rate	55.1%	49.3%	55.1%	37.7%
Area under the ROC curve	76.7%	77.7%	95.9%	92.4%
Area under the PRC curve	96.7%	95.7%	99.6%	99.3%

Table 4 shows the performance of the four different algorithms to predict vaccination tendency for children. In this forecasting evaluation process, the child-vaccination dataset contains missing observations that are more than 24 times as those in the adults' vaccination dataset. As a result, the child vaccination data only has 688 records with 65 attributes. Due to a higher data sparsity and a smaller number of records, predicting the vaccination hesitancy for children generates less accuracy and precision, as shown in Table 4. Similarly, the RF and ANN models perform better than the other two models and the RF model outperforms the three other models in most of the metrics such as Accuracy, Recall, F-Measure, MCC, True positive rate, False positive rate, Area under the ROC curve and Area under the PRC curve.

Table 4 Model's Performance to Predict Children Vaccination Tendency

	LR	DT	RF	ANN
Accuracy	82.7%	81.5%	86.5%	83.0%
Precision	88.8%	87.4%	86.2%	87.9%
Recall	90.2%	90.4%	99.3%	91.8%
F-Measure	89.5%	88.9%	92.3%	89.8%
MCC	40.6%	34.6%	46.8%	38.8%
True positive rate	90.2%	90.4%	99.3%	91.9%
False positive rate	50.8%	57.9%	70.6%	56.3%
Area under the ROC curve	79.1%	65.0%	87.5%	82.6%
Area under the PRC curve	91.6%	84.7%	97.0%	94.5%

5. **Feature ranking**

Table 5 Feature importance to predict self-vaccination for different machine learning models

Attribute	**ANN**	**LR**	**DT**	**RF**
Vaccine attitude: social responsibility	18.2	30.33	-4.04	78.87
Test_difficulty	-4.04	50.56	2.02	68.76
Test reason: recommended_by_health	-8.09	61.68	0	38.42
Overconfidence	-8.09	4.04	-4.04	24
Age	4.04	26.29	0	20.22
Vaccine attitude: more relaxed with vaccine	34.38	22.24	66.73	6.07
Test location: clinic	-6.07	24.27	4.04	6.07
Vaccine reason: paid_leave	-10.11	71.79	2.02	4.04
Test reason: contacted_by_health	-22.24	92.01	6.07	2.02
Test barrier: time	-6.07	36.4	4.04	-8.09
Test location: drive_thru	4.04	8.09	9.1	-10.11
Test reason: work_required	-2.02	26.29	8.09	-12.13
flushot	18.2	86.96	-4.04	-12.13
Vaccine reason:close_by	32.36	46.51	14.16	-20.22
Information resource: pharmacy	-10.11	71.79	9.1	-24.27

Table 6 shows the set of top ranked feature importance for the four models to predict children vaccination intention. Although the results still present some variations among feature importance obtained by the different Machine Learning models, the rankings between the DF and RF models are closer compared with those of the other two models. However, the rankings from the best two performing models ANN and RF present little correlation among themselves.

Table 6 Feature importance to predict children's vaccination for different machine learning models

Attribute	**ANN**	**LR**	**DT**	**RF**
Information resource: employer	0.87	-3.34	6.25	26.89
Information source: pharmacy	4.65	-0.58	5.52	23.26
Vaccination reason: no_side_effect	0.00	0.29	0.29	21.80
family_composition	1.45	-2.91	0.00	20.93
Vaccination reason: paid_leave	0.29	7.56	6.98	20.64
race	1.16	-1.74	8.43	20.06
Vaccination reason: transportation	2.62	-5.52	5.52	17.73
Vaccination attitude: social_responsibility	0.87	0.00	7.70	17.73
Vaccination reason: others getting it	0.29	-4.07	1.45	16.57
Education	2.62	-2.33	0.29	14.83
Overconfidence	-7.849	-0.872	-5.523	14
Information resource: insurance_company	2.62	1.16	5.52	12.79
Vaccination reason: required	2.04	0.00	0.87	12.79

Test barrier: insurance	2.04	-3.34	5.52	11.63
flushot	5.23	-0.58	4.80	10.76

Since the RF model performs the best in most of the measures, we use RF to find the most relevant features to predict vaccination tendency for both adults themselves and their children. In the following figure, we list the top ranked features to predict vaccination tendency by the RF model.

Figure 1 Feature importance to predict self-vaccination intention by the RF model

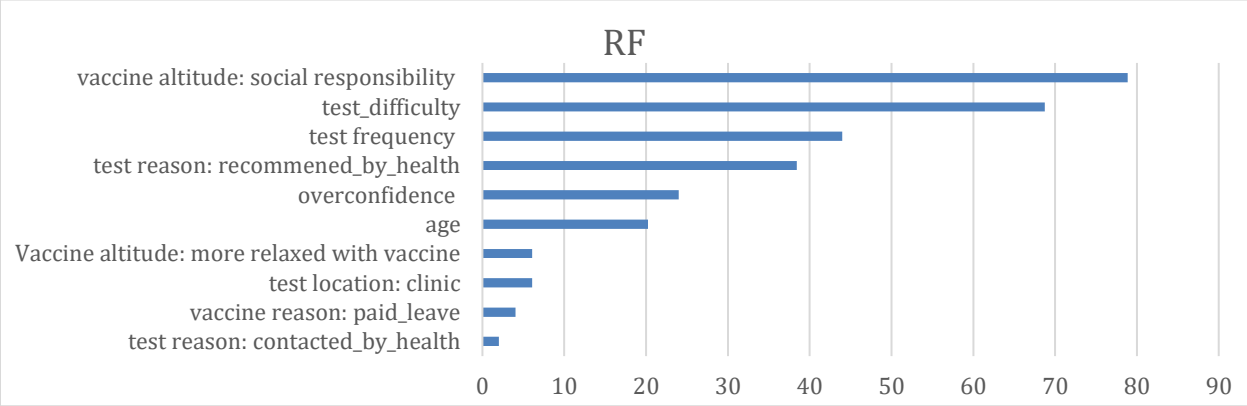

Figure 1 reports the top 10 most relevant features to predict self-vaccination tendency. Similar to the results presented by the linear regression model, the belief of vaccination as a social responsibility and test frequency are the important features to predict an adult's vaccination tendency. Other relevant features that are also significant according to linear regression model include 'paid leave' and 'vaccine attitude: more relaxed with vaccine.'

Figure 2 Feature importance to predict children vaccination intention by RF model

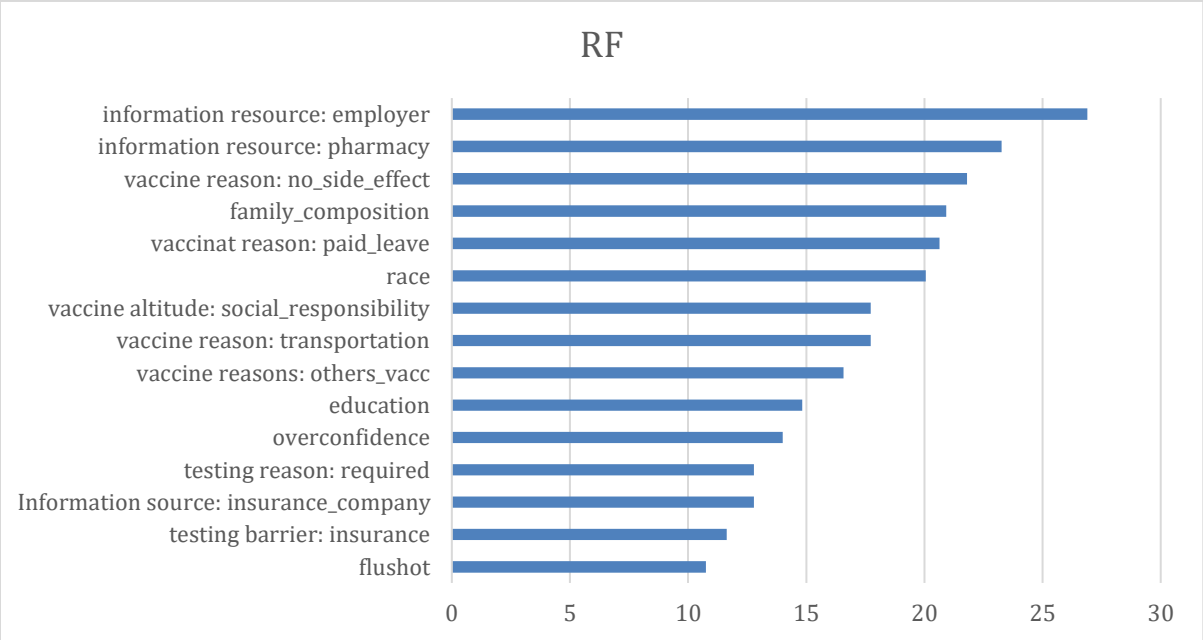

Figure 2 shows the top 15 relevant features to predict child vaccination tendency. The common important features shared with self-vaccination are 'paid leave', 'social responsibility', 'overconfidence'. Once again, the RF model supports the results from the multiple regression model in which 'paid leave' and 'social responsibility' are the important features to predict children vaccination intention. Compared with self-vaccination intention, more features from the information resources and reasons to get vaccinated became important. These results imply that parents make vaccination decision for their children based on a set of slightly different factors. The sources of information where people learn about COVID-19 vaccines such as their employer, pharmacy and insurance company have a strong influence in their decision making for vaccinating their children. One possible explanation is that people are more cautious in getting their children vaccinated so that they become more aware of and more selective about the information source when educating themselves about the COVID-19 vaccines. Combining this with the previous result that the percentage of people who are willing to vaccinate their children is higher than the percentage to vaccinate themselves, certain information sources to learn about the COVID-19 vaccines help people perceive the advantages of receiving the vaccines more than the vaccine's potential adverse side effects.

6. Practical implication

To be added

7. Discussion

This study examines the factors that influence people's decision to be vaccinated and their intentions to vaccinate their children. Findings suggest that testing frequency, social responsibility and paid leave are strong predictors of people's vaccination decisions for themselves and their children. One possible explanation for the positive correlation between testing frequency and vaccination is that frequent required testing (due to work, travel, etc.) increases the cost of not getting vaccinated. Therefore, people prompt to get vaccinated to avoid the psychological loss and opportunity costs, such as loss in time, potential earning, chances of being infected by the coronavirus, getting seriously ill, etc. Another possible explanation could be that there are unobservable characteristics that explain both testing and vaccination behavior, such as overconfident bias, conformity of social norms, and effect heuristic.[6] Future research should further examine the relationship between testing and vaccination behaviors strategies through the lens of behavioral economics. Similarly, people more likely get vaccinated for themselves and their children if the costs associated with getting ill, loss in time and work can be partially compensated with paid leave.

Based on the findings in the study, we provide the following policy suggestions to increase adherence to COVID-19 vaccination: allocate more resources to communities with a lower testing rate; leverage existing testing sites to educate and promote the social responsibility of vaccination; incentivize employers to offer paid leave to their workers; reduce the costs of

[6] Please see the extensive discussion on applying behavioral economics to understand COVID-19 behaviors by Soofi, Najafi and Karami-Matin (2020).

vaccination by streamlining the procedure to mitigate vaccination difficulties; moreover, invest in public health information structure to make the learning more accessible.

References

Al-Amer, R., Maneze, D., Everett, B., Montayre, J., Villarosa, A. R., Dwekat, E., & Salamonson, Y. (2022). COVID-19 vaccination intention in the first year of the pandemic: A systematic review. Journal of clinical nursing, 31(1-2), 62-86.

Albrecht, D. (2022). Vaccination, politics and COVID-19 impacts. BMC Public Health 22, 96. https://doi.org/10.1186/s12889-021-12432-x.

Ali, K. F., Whitebridge, S., Jamal, M. H., Alsafy, M., & Atkin, S. L. (2020) Perceptions, knowledge, and behaviors related to COVID-19 among social media users: cross-sectional study. J Med Internet Res 22:e19913.

Bagasra, A. B., Doan, S. & Allen, C.T. (2021). Racial differences in institutional trust and COVID-19 vaccine hesitancy and refusal. BMC Public Health 21, 2104. https://doi.org/10.1186/s12889-021-12195-5

Bailey, S. C., Serper, M., Opsasnick, L., Persell, S. D., O'Conor, R., Curtis, L. M., ... & Wolf, M. S. (2020). Changes in COVID-19 knowledge, beliefs, behaviors, and preparedness among high-risk adults from the onset to the acceleration phase of the US outbreak. *Journal of general internal medicine*, *35*(11), 3285-3292.

Banik, R., Islam, M., Pranta, M.U.R. et al. Understanding the determinants of COVID-19 vaccination intention and willingness to pay: findings from a population-based survey in Bangladesh. *BMC Infect Dis* 21, 892 (2021). https://doi.org/10.1186/s12879-021-06406-y

Barrios, J. M., and Hochberg, Y. (2020). Risk Perception through the Lens of Politics in the Time of the COVID-19 Pandemic. National Bureau of Economic Research. Cambridge, United states.

Barry, V., Dasgupta, S., Weller, D., et al. (2021). Patterns in COVID-19 Vaccination Coverage, by Social Vulnerability and Urbanicity, Morbid Mortal Wkly Rep; 70(22): 818–24.

Brinson, N. H. (2022). Resistance to Persuasion: Examining the Influence of Political Ideology on COVID-19 Vaccine Uptake Hesitancy. J Frontiers in Communication. https://www.frontiersin.org/article/10.3389/fcomm.2021.760847

Carrieri, V., Lagravinese, R., & Resce, G. (2021). Predicting vaccine hesitancy from area-level indicators: A machine learning approach. *Health Economics*, *30*(12), 3248-3256.

Cheng, K.J.G., Sun, Y. and Monnat, S.M. (2020), COVID-19 Death Rates Are Higher in Rural Counties With Larger Shares of Blacks and Hispanics. *The Journal of Rural Health*, 36: 602-608. https://doi.org/10.1111/jrh.12511

Chu, J., Pink, S. L., & Willer, R. (2021). Religious identity cues increase vaccination intentions and trust in medical experts among American Christians, Proceedings of the National Academy of Sciences, 118 (49) e2106481118, https://www.pnas.org/doi/abs/10.1073/pnas.2106481118

Crozier, J., Christensen, N., Li, P., Stanley, G., Clark, D. S., & Selleck, C. (2021). Population Health Management. http://doi.org/10.1089/pop.2021.0216

Detoc, M.; Bruel, S.; Frappe, P.; Tardy, B.; Botelho-Nevers, E.; Gagneux-Brunon, A. Intention to participate in a COVID-19 vaccine clinical trial and to get vaccinated against COVID-19 in France during the pandemic. *Vaccine* 2020, 38, 7002–7006.

Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

Fernandes, N., Costa, D., Costa, D., Keating, J., & Arantes, J. (2021). Predicting COVID-19 vaccination intention: the determinants of vaccine hesitancy. *Vaccines*, *9*(10), 1161.

Fridman A, Gershon R, Gneezy A. (2021). COVID-19 and vaccine hesitancy: A longitudinal study. Capraro V, editor. PLoS One. 16(4):e0250123.

Gao, G., & Li, D. (2022). Knowledge, overconfidence, and behavior in COVID-19: results from an online survey. *Econ Bull*, *42*(1), 215-223.

Hawryluk, M. (2021). COVID-19: rural Americans in pharmacy deserts hurting for Covid vaccines. San Francisco, CA: Kaiser Family Foundation. https://khn.org/news/article/rural-america-pharmacy-deserts-hurting-for-covid-vaccine-access/

Huynh, G., Tran, T. T., Nguyen, H. T. N., & Pham, L. A. (2021). COVID-19 vaccination intention among healthcare workers in Vietnam. *Asian Pacific Journal of Tropical Medicine*, *14*(4), 159.

Iwendi, C., Bashir, A. K., Peshkar, A., Sujatha, R., Chatterjee, J. M., Pasupuleti, S., ... & Jo, O. (2020). COVID-19 patient health prediction using boosted random forest algorithm. *Frontiers in public health*, 8, 357.

Kenji Kira, Larry A. Rendell: A Practical Approach to Feature Selection. In: Ninth International Workshop on Machine Learning, 249-256, 1992.

Kerr, J. R., Freeman, A., Marteau, T. M., & van der Linden, S. (2021). Effect of Information about COVID-19 Vaccine Effectiveness and Side Effects on Behavioural Intentions: Two Online Experiments. Vaccines, 9(4), 379. https://doi.org/10.3390/vaccines9040379

Ku L. (2021). The Association of Social Factors and Health Insurance Coverage with COVID-19 Vaccinations and Hesitancy. J. Gen Intern Med., 37(2):409-414. doi: 10.1007/s11606-021-07213-6.

Latkin, C. A., Dayton, L,, Yi, G., Colon, B., & Kong, X. (2021) Mask usage, social distancing, racial, and gender correlates of COVID-19 vaccine intentions among adults in the US. PLOS ONE 16(2): e0246970. https://doi.org/10.1371/journal.pone.0246970

Lee, M., Kang, B. A., & You, M. (2021). Knowledge, attitudes, and practices (KAP) toward COVID-19: a cross-sectional study in South Korea. *BMC public health*, *21*(1), 1-10.

Lewis, J. R. (2020). What is driving the decline in people's willingness to take the COVID-19 vaccine in the United States? JAMA Health Forum. 1:e201393.

Lin, Y., Hu, Z., Zhao, Q., Alias, H., Danaee, M., & Wong, L. P. (2020) Understanding COVID-19 vaccine demand and hesitancy: a nationwide online survey in China. PLOS Neglect Trop D 14:e0008961. https://doi.org/10.1371/journal.pntd.0008961

Lin, C., Tu, P., Beitsch, L.M. (2021). Confidence and receptivity for covid-19 vaccines: A rapid systematic review. Vaccines. 9(16). doi: 10.3390/vaccines9010016.

Mackey, K., Ayers, C. K., Kondo, K. K., Saha, S., Advani, S. M., Young, S., et al. (2021). Racial and ethnic disparities in COVID-19–related infections, hospitalizations, and deaths: a systematic review. Ann Intern Med. 174(3):362–73. https://doi.org/10.7326/M20-6306.

Mellado, B., Wu, J., Kong, J. D., Bragazzi, N. L., Asgary, A., Kawonga, M., ... & Orbinski, J. (2021). Leveraging artificial intelligence and big data to optimize COVID-19 clinical public health and vaccination roll-out strategies in Africa. *International Journal of Environmental Research and Public Health*, *18*(15), 7890.

Muhammad, L. J., Islam, M. M., Usman, S. S., & Ayon, S. I. (2020). Predictive data mining models for novel coronavirus (COVID-19) infected patients' recovery. *SN Computer Science*, 1(4), 1-7.

Murthy, B. P., Sterrett, N., Daniel Weller, D., et al. (2021). Disparities in COVID-19 Vaccination Coverage Between Urban and Rural Counties — United States, December 14, 2020–April 10, 2021. MMWR Morb Mortal Wkly Rep 2021;70:759–764. DOI: http://dx.doi.org/10.15585/mmwr.mm7020e3

Olagoke, A. A., Olagoke, O. O., & Hughes, A. M. (2021). Intention to Vaccinate Against the Novel 2019 Coronavirus Disease: The Role of Health Locus of Control and Religiosity. Journal of religion and health, 60(1), 65–80. https://doi.org/10.1007/s10943-020-01090-9

Ong, E., Wong, M. U., Huffman, A., & He, Y. (2020). COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning. *Frontiers in immunology*, 11, 1581.

Paul, E., Steptoe, A., & Fancourt, D. (2021). Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications, The Lancet Regional Health - Europe, Volume 1, 100012, ISSN 2666-7762, https://doi.org/10.1016/j.lanepe.2020.100012.

Petersen, M.B., Bor, A., Jørgensen, F.J., & Lindholt, M.F. (2020). Transparency Is Necessary but Not Sufficient to Reduce Skepticism About a COVID-19 Vaccine. PsyArXiv. doi: 10.31234/OSF.IO/VX84N

Pink, S. L., Chu, J., Druckman, J. N., Rand, D. G., & Willer, R. (2021). Elite party cues increase vaccination intentions among Republicans. Proc. Natl. Acad. Sci. U.S.A. 118, e2106559118

Rahman, M. M., Chisty, M. A., Alam, M. A., Sakib, M. S., Quader, M. A., Shobuj, I. A., ... & Rahman, F. (2022). Knowledge, attitude, and hesitancy towards COVID-19 vaccine among university students of Bangladesh. *PloS one*, *17*(6), e0270684.

Ritonga, M., Al Ihsan, M. A., Anjar, A., & Rambe, F. H. (2021, February). Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm. In IOP Conference Series: *Materials Science and Engineering*, 1088, (1), p. 012045. IOP Publishing.

Robinson, E., Jones, A., & Daly, M. (2020). International estimates of intended uptake and refusal of COVID-19 vaccines: a rapid systematic review and meta-analysis of large nationally representative samples. medRxiv. https://doi.org/10.1101/2020.12.01.20241729

Ryan J. Urbanowicz, Melissa Meeker, William La Cava, Randal S. Olson, Jason H. Moore, Relief-based feature selection: Introduction and review, *Journal of Biomedical Informatics*, Volume 85, 2018, Pages 189-203.

Sherman, S.M.; Smith, L.E.; Sim, J.; Amlôt, R.; Cutts, M.; Dasch, H.; Rubin, G.J.; Sevdalis, N. COVID-19 vaccination intention in the UK: Results from the COVID-19 vaccination acceptability study (CoVAccS), a nationally representative cross-sectional survey. *Hum. Vaccin. Immunother*. 2021, 17, 1612–1621.

Solís Arce, J. S., Warren, S. S., Meriggi, N. F. et al. (2021). COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. Nat Med 27, 1385–1394. https://doi.org/10.1038/s41591-021-01454-y

Sood, L. and Sood, V. (2021), Being African American and Rural: A Double Jeopardy From COVID-19. The Journal of Rural Health, 37: 217-221. https://doi.org/10.1111/jrh.12459)

Soofi, M., Najafi, F., & Karami-Matin, B. (2020). Using insights from behavioral economics to mitigate the spread of COVID-19. *Applied health economics and health policy*, *18*(3), 345-350.

Thunström, L., Ashworth, M., Finnoff, D. et al. (2021). Hesitancy Toward a COVID-19 Vaccine. EcoHealth 18, 44–60. https://doi.org/10.1007/s10393-021-01524-0

Viswanath, K., Bekalu, M., Dhawan, D., Pinnamaneni, R., Lang, J., & McLoud, R. (2021). Individual and social determinants of COVID-19 vaccine uptake. BMC Public Health. 21(1):818. https://doi.org/10.1186/s12889-021-10862-1.

Wang, J., Zhu, H., Lai, X., Zhang, H., Huang, Y., Feng, H., ... & Fang, H. (2022). From COVID-19 vaccination intention to actual vaccine uptake: A longitudinal study among Chinese adults after six months of a national vaccination campaign. *Expert review of vaccines*, *21*(3), 385-395.

Wong, L.P., Alias, H., Danaee, M. et al. (2021). COVID-19 vaccination intention and vaccine characteristics influencing vaccination acceptance: a global survey of 17 countries. Infect Dis Poverty 10, 122. https://doi.org/10.1186/s40249-021-00900-w

Zhang, X., Thompson, P., Ras, Z.W., and Jastreboff, P. (2011). Mining Tinnitus Data Based on Clustering and New Temporal Features, in Learning Structure and Schemas from Documents, M. Biba, F. Xhafa (Eds.), *Studies in Computational Intelligence*, Vol. 375, Springer, 2011, 227-246.

Zintel, S., Flock, C., Arbogast, A.L. et al. Gender differences in the intention to get vaccinated against COVID-19: a systematic review and meta-analysis. J Public Health (Berl.) (2022). https://doi.org/10.1007/s10389-021-01677-w

Appendix: Survey

Start of Block: Block 1

We are requesting your participation in a research study entitled: Building Resiliency and Vital Equity (BRAVE) project. This is a NIH funded project with the University of North Carolina Pembroke, North Carolina Central University, and the Lumbee Tribe of NC. The purpose of this study is to understand people's perceptions/beliefs about covid-19 testing and vaccination, and to help the communities fight COVID-19. You must be 18 years or older to participate in this survey. You will answer demographic questions, (including, age, race, gender, education, income, etc.) and questions related to your attitude, behaviors, and knowledge of COVID-19. Participation in this survey is completely voluntary, and you may withdraw from the survey at any time. All of your answers will remain confidential and will be used for research purpose only. The survey has about 40 questions and takes about 15 minutes. Whiling completing the survey, you can skip any question or withdraw from the study anytime. If you withdraw from it, your data may still be included in the analyses. If you decide to withdraw, please inform the research of it. You will receive $10 Walmart gift card for filling out the survey even if you decide to withdraw from the survey. You can choose to have this gift card mailed or you can pick it up in person (more information provided at the end of the survey). You will receive $10 Walmart gift card within a month of completing the surveyI have read the above information and agree to proceed.

○ I do not want to participate in this survey.

End of Block: Block 1

Start of Block: Default Question Block

*

Age

_____

Race and ethnicity? (Check all that apply).

- [ ] Asian
- [ ] Alaska Native
- [ ] Black or African American
- [ ] Hispanic or Latino
- [ ] Native American or American Indian
- [ ] Native Hawaiian or other Pacific Islanders
- [ ] White
- [ ] Prefer not to say

Highest level of education

- ○ Did not complete high school
- ○ High school diploma or GED
- ○ Some college or trade school
- ○ College - bachelor's degree
- ○ Graduate or professional degree

What best describes your family at home?

- ○ Just me
- ○ Living with spouse, no kids
- ○ Family including kids
- ○ Family with 3 generations (parents, children, grandchildren)
- ○ Family with 4 generations
- ○ None of these

What is your employment status?

- ○ Working full-time
- ○ Working part-time
- ○ Only temporarily laid off
- ○ Sick leave or maternity leave
- ○ Looking for work, unemployed
- ○ Retired
- ○ Disabled permanently or temporarily
- ○ Keeping house
- ○ Student
- ○ Prefer not to answer
- ○ Don't know

What was your household’s total income last year before taxes? Include everyone in your household

- ○ Less than $15,000
- ○ $15,000-19,999
- ○ $20,000-24,999
- ○ $25,000-34,999
- ○ $35,000-49,999
- ○ $50,000-74,999
- ○ $75,000-99,999
- ○ $100,000 or above
- ○ Prefer not to answer

Gender you identify with?

- ○ Male
- ○ Female
- ○ Nonbinary
- ○ Transgender
- ○ Something else
- ○ Prefer not to answer

What is your residential zip code?

_____

What health insurance do you have?

- ○ I do NOT have health insurance
- ○ Private (purchased directly or through employment)
- ○ Public (Medicare, Medicaid, Tricare)
- ○ Do not know
- ○ Prefer not to answer

Since the pandemic began, how may times (if any) did you get tested for COVID? (your best guess if tested regularly)

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9
- ○ 10 or more

What were the reasons that you got tested? (check all that apply)

- [ ] Wanted to know
- [ ] Had symptoms
- [ ] Exposed to someone who had COVID or symptoms
- [ ] Required for my job
- [ ] Contacted by the health department
- [ ] Recommended by health care provider
- [ ] Not applicable

Where did you get tested? (check all that apply)

- [ ] Clinic
- [ ] Drive-thru
- [ ] Home, used home testing kit
- [ ] Mobile unit
- [ ] Lab
- [ ] Community location or event (such as churches, community centers)
- [ ] Not applicable

Page Break

Since the pandemic began, was there ever a time when you wanted to get a COVID test, but could not?

- ○ Yes
- ○ No

At that time, what prevented you from getting a COVID test? (check all that apply, even if tests are easier to get now)

- ○ No place (at that time) to get tested in the community
- ○ No transportation
- ○ No time
- ○ No insurance

If you wanted to get a COVID test today, do you know where or how to get it?

- ○ Yes
- ○ No

How difficult would it be?

- ○ Extremely difficult
- ○ Difficult
- ○ Not difficult

Have you had COVID? (Have you tested positive for COVID or told by a health care provider that you had it)

- ○ Yes
- ○ No

Page Break

Which, if any, of the following sources have you used to find information about getting a COVID vaccine? (Check all that apply).

- [ ] The "Your Spot Your Shot" Website
- [ ] My health department
- [ ] My health insurance company
- [ ] My medical provider
- [ ] My pharmacy
- [ ] My employer
- [ ] Friends/family
- [ ] Community groups
- [ ] Faith leaders

Page Break

Please answer the following questions (True or False):

	True	False	Do not know
The only people who can get severely ill from COVID-19 are the elderly and people with certain medical conditions.	○	○	○
People who have COVID-19 and no symptoms can still spread it to others.	○	○	○
If someone tests negative for COVID-19, this person can still get infected in the future.	○	○	○
COVID-19 vaccines help keep people from getting seriously ill even if they do get infected with COVID-19.	○	○	○
Fully vaccinated people do not need to be tested after being exposed to someone with COVID-19.	○	○	○

How confident are you about your answers in the chart above?

	Not confident at all	Slightly confident	Somewhat confident	Fairly confident	Extremely confident
Choose	○	○	○	○	○

How confident do you think an average person can answer these questions correctly?

	Not confident at all	Slightly confident	Somewhat confident	Fairly confident	Extremely confident
Choose	○	○	○	○	○

Page Break

How easy or difficult has it been for you to:

	Very easy	Somewhat easy	Somewhat difficult	Very difficult	N/A
Get information about COVID vaccines.	○	○	○	○	○
Sign up for a COVID vaccine.	○	○	○	○	○
Find a convenient time and place to be vaccinated.	○	○	○	○	○

Page Break

Which one of the following best describes your COVID-19 vaccination status? (Select one.)

- ○ I plan to get vaccinated.
- ○ I have received the first shot for the 2-shot vaccine.
- ○ I am fully vaccinated.
- ○ I have no plans to get vaccinated.

Please tell us more about your vaccine plans.

- ○ I have an appointment
- ○ I am on a waiting list but do not have an appointment yet.
- ○ I have tried to get an appointment but I am not able to get one.
- ○ I am planning to get vaccinated but have not made an appointment yet.

Please tell us more about getting the second shot.

- ○ I have an appointment.
- ○ I need to make an appointment to get the 2nd shot.
- ○ I will NOT get the 2nd shot because of my reaction to the 1st shot.
- ○ I will NOT get the 2nd shot because one shot is sufficient enough to protect me.
- ○ I will NOT get the 2nd shot because of herd immunity.

Please tell us why you do not plan to get vaccinated. (Select all apply.)

- [ ] allergic to vaccines
- [ ] don't like needles
- [ ] concerned about getting COVID-19.
- [ ] concerned about side effects from the vaccine.
- [ ] don't think vaccines work very well
- [ ] don't trust the vaccine
- [ ] don't believe the COVID-19 pandemic is as bad as some people say it is.
- [ ] don't want to pay for it.
- [ ] don't know enough about how COVID-19 vaccine works.
- [ ] My health condition (other than allergies) does not allow me to get vaccinated.

What are the most important things to you regarding COVID vaccination? (Check all that apply)

- [ ] Vaccination offered close to where I live.
- [ ] Employer gives me paid time off work.
- [ ] Employer asks or requires me to get vaccinated.
- [ ] Able to take time off after vaccination in case of side effects (pain, swelling, fever, chills, tiredness etc.).
- [ ] My nurse or doctor told me to get it.
- [ ] Hearing from people that I know who have gotten it.
- [ ] Seeing people like me get the vaccine.
- [ ] Seeing people who didn't get side effects.
- [ ] Have transportation to the vaccination location.
- [ ] Able to get childcare or bring my children with you.
- [ ] Vaccine given at a place where I am comfortable such as a church, school, or community center.

Page Break

How much do you agree or disagree with each statement?

	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I am or will be more relaxed once I can get the vaccine.	○	○	○	○	○
I may not wear a mask after fully vaccinated.	○	○	○	○	○
Once the COVID vaccine is given to most of the population, our lives will finally return to normal again.	○	○	○	○	○
It is my responsibility to get vaccinated so that our whole community is protected.	○	○	○	○	○

Page Break

Do you plan to or have you already gotten your child/children vaccinated?
Yes, I plan to or have already had my child/children get vaccinated.

- ○ No, I do not plan to get my child/children vaccinated.
- ○ Not applicable – I don’t have children that I take care of.

In the past 12 months, did you get a flu shot?

- ○ Yes
- ○ No
- ○ Don't remember
- ○ Never get the flu shot

In the past 12 months, how many children in your care have gotten flu shots?

- ○ All or some of them
- ○ None of them
- ○ Don't remember
- ○ Not applicable – I don’t have children that I take care of

---

Page Break

How would you describe your life now compared to before the COVID-19 pandemic?

	Worse off	Better off	About the same
Financially, I am _____	○	○	○
Physically, I am _____	○	○	○
Emotionally, I am _____	○	○	○

Page Break

How confident are you that COVID-19 pandemic will end during 2021?

	Choose One
Very Confident	○
Fairly confident	○
Somewhat confident	○
Slightly confident	○
Not confident at all	○

Page Break

Thank you for participating in this survey. We appreciate your time. You will receive your payment in the form of a $10.00 Walmart gift card within a month. Please let us know how you would like to receive this gift card by choosing from one of the options below.

- ○ I would like this gift card to be mailed to me. Please provide your mailing address.
  _____
- ○ I would like to pick up this gift card in person.

Page Break

Please confirm your mailing address you provided.

_____

Page Break

You can pick up your gift card from Jessica Jones at Lumbee Tribe of NC Veteran's Office located at 204 W 3rd Street, Pembroke (910) 405-0501, jljones@lumbeetribe.com

End of Block: Default Question Block