

## Using Mobile Phone-based Text Message to Recruit Representative Samples: Assessment of a Cross-Sectional Survey about the COVID-19 Vaccine Hesitation.

Sartorao Filho, C.I.<sup>1,2,3</sup>, Sartorao Neto, C.I.<sup>4</sup>, Sartorao, A.L.V.<sup>1</sup>, Terribile, D.C.<sup>1</sup>, Mello, R.<sup>2</sup>, Mello, B.B.<sup>1</sup>, Zoqui, M.C.<sup>1</sup> Duarte, D.O.<sup>1</sup> Cachoni, L.E.G.,<sup>1</sup> Bisseto, V.C.Q.<sup>1</sup>, Ribeiro, E.A.C.<sup>1</sup>

1 Educational Foundation of the Municipality of Assis, Faculty of Medicine;

2 Sao Paulo Medical Association;

3 Sao Paulo State University - Botucatu Medical School;

4 University of Sao Paulo

Corresponding Author:

Carlos Izaias Sartorão Filho [carlos.sartorao@unesp.br](mailto:carlos.sartorao@unesp.br)

Benedito Spinard street, 1440 Assis-SP – Brazil. ZIP 19815-110

Phone: +5518 33242122

### Abstract

**Background:** Limited research has examined mobile phone-based platforms for survey recruitment, especially during the COVID-19 pandemic in Brazil. Our objective was to investigate the feasibility and representativeness of mobile phone-based advertisement during a preliminary study about COVID-19 vaccine hesitation in Brazil. Moreover, we evaluate whether the older population can be reached through mobile phone-based platforms of the survey. **Methods:** We conducted a study in December 2021 based on a preliminary survey about the COVID-19 vaccine hesitation in Assis, Brazil, Sao Paulo state. From a list of the adult population hesitant for the second dose of the COVID-19 vaccine, we sent a mobile phone-based advertisement inviting the participants to answer the survey for one week. The respondent's data were collected in a Google form platform. The comparison between the target population and the respondents was made using the Chi-squared test and the Welch's test, using a P-value of .05 as significant. **Results:** The response rate was 9.99% after one week. The mean age of the respondent group was 33.97 (SD 14.99) and 35.05 (SD 14.19) of the population, with a *P-value* of .192 and a Cohen's d coefficient of 0.0754, corresponding to a small effect size between groups. We demonstrate that the mobile phone-based survey is a feasible and representative strategy during the pandemics in Brazil. Moreover, the older population respondent was representative. **Conclusion:** We achieved a representative sample of respondents using the mobile phone-based survey in Brazil. Furthermore, it was representative in all sociodemographic and health characteristics assessed. Finally, these findings suggest the method is a highly feasible and economical means of recruiting for survey research.

**Keywords:** Survey and Questionnaires; Demography; Cell Phone; COVID-19; Vaccination Hesitancy.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## Introduction

Population-based survey research is useful for recruitment and aims for collected samples to be representative of the target population. Otherwise, traditional survey research can be limited by the higher costs, low response rates, and demands more personnel and time be made. (1) In addition, landline telephones are recognized in disuse, limiting their ability for recruitment. Limited research has examined mobile phone-based platforms for survey recruitment, especially in Brazil. (2) Besides, it remains unclear whether the method can recruit a representative sample of older adults (3,4). Therefore, we aimed to investigate whether mobile phone-based text messages can be a feasible, timely, economical, and with minimal human resource commitment means of recruiting a representative sample of adults for a health research survey based on the hesitance for the COVID-19 vaccine in Brazil.

## Methods

At the Educational Foundation of the Municipality of Assis, we conducted a study based on a preliminary survey about the reasons for hesitating or refusing the COVID-19 vaccine in Assis, Brazil, São Paulo state. The Institution Ethical Approval was obtained; CAAE: 51936621.8.0000.8547/Statement: 5.135.036. Consent was obtained before individuals could begin the survey, and if it was not provided, participants were redirected away from the survey. The survey itself was anonymous, and all efforts were made to maintain the confidentiality of individuals who participated. The online text message was used only to advertise and direct interested individuals to the survey link due to COVID-19 vaccine hesitation, hosted with Google Forms. Data were securely kept on password-protected computers and cloud storage accounts. To ensure transparency in the study design and recruitment process, we followed the guidelines for reporting results of internet surveys, describing the checklist details throughout our methods. (5) No incentives were offered for participation. The central theme of our survey was the second dose of COVID-19 vaccine hesitation, focusing on health behaviors, the reasons for, and the self-reported attitude to complete or not the proposed vaccine schedule. A mobile phone-based invitation was sent to all adults from a list provided by the official health authorities, who were delayed more than 30 days for the second dose of the vaccine. Google Forms was used to host the open survey and allow for automatic capture of responses into a spreadsheet. We did not use item randomization or adaptive questioning. There were ten items and four screens, one for the invitation, one for the introduction and consent, one for the survey, and one to thank participants and offer the opportunity to contact us or provide feedback. No completeness check was used, but participants could return to previous pages and change responses. The questions were about the sociodemographic characteristics, the reasons for vaccine hesitation, and the different attitudes toward completing the vaccine proposed scheme.

The official municipality health department provided a list of participants with at least 30 days of delay to the second dose of the COVID-19 vaccine in Assis. The period of observation was from January 01 to November 03, 2021. The health government institution previously agreed to participate in the study. Therefore, we used the spreadsheet containing the name, birth date, gender, address, mobile phone number, date of the first dose, vaccine brand name, the presumed date of the second dose.

Advertisements ran twice a week for each participant's mobile phone number, during December 10-17, 2018, until the cutoff date for the recruitment. We eligible all the adult participants listed in Assis's municipality as having the hesitation or refusal for the second vaccine dose more than one month of delay. Invalid, missing data, or inexistent cell phone numbers were excluded.

Data analyses comprised the following three major aspects: data cleaning and checking, descriptive analysis of online text messages ad metrics and costs to determine cost per recruit, and descriptive analyses to provide an overview of our sample and univariate analysis to examine whether the distributions of the sample's sociodemographic and selected health characteristics were consistent with the target population. All statistical analyses were conducted using Microsoft Excel Version 16.

Concerning response rates, the number of respondents was determined with Google Forms. Unfortunately, Google Forms does not track the number of surveys started—only the number of surveys submitted—and there is no way to prevent or identify when there are multiple entries from the same individual unless they were precise duplicates.

Feasibility was assessed as the ease of conducting research and time commitment and objectively by the costs per recruit. Recruitment costs included advertising costs and administrative costs during the entire recruitment period.

Representativeness of the mobile text message sample was assessed by comparing sociodemographic characteristics of participants respondents with the underlying population, obtained from the census provided by the Health Municipality. These characteristics were age, gender, self-reported comorbidities, type of the first dose vaccine, and the number of delayed days since the recommended second dose date.

We performed Welch's procedure to compare the two groups (adult population delayed for the COVID-19 vaccine and group of respondents to the mobile phone-based survey), measurement means, and standard deviation. The alpha level was 0.05, two-tails, and the control event rate was 20%. Using an Excel-based calculator, we calculate the mean difference confidence intervals, effect sizes, and Levene's test of inequality of variance. (6)

The Chi-squared goodness of fit test was used to test if sample data fits a distribution from the population. For statistical analysis,  $P < .05$  was considered to indicate rejection of the null hypothesis at the significance level.05. Likewise,  $P > .05$  indicates a failure to reject the null hypothesis at the significance

level.05, meaning we can assume the sample distribution is consistent or representative with the census distribution.

The magnitude of difference between the respondents' distribution and the population, for effect size, was calculated and interpreted as per Cohen (7), adopting values of 0.20, 0.50, and 0.80 correspond to small, medium, and large effect sizes. We considered a character to represent the population if the Cohen's  $d < 0.20$  (small effect size).

## Results

There were 4667 vaccine-hesitant adults on the list given by the health authorities from the municipality, delayed for more than 30 days, a mean of 91.39 days (+/- 37.68). In addition, 1120 were excluded due to missing, invalid, or inconsistent mobile phone numbers. Thus, we sent 3547 mobile phone-based text messages. Finally, after one week, we obtained 355 respondents, a rate response of 9.99% after one week of surveillance. Table 1 describes the respondent's characteristics.

**Table 1. Respondents Characteristics (N=355)**

Characteristic	Data*
<b>Age, Years</b>	
Mean (SD)	33.98 (14.99)
Range	18-88
<b>Age Group, years</b>	
18-29	179 (50.4)
30-44	96 (27.0)
45-59	53 (14.9)
≥60	27 (7.6)
<b>Gender</b>	
Female	153 (43.1)
Male	202 (56.9)
<b>Educational attainment</b>	
College graduate or above	80 (22.5)
<b>Have you had COVID-19?</b>	
Yes (confirmed or clinically suspected)	37 (10.4)
<b>Self-rated overall health</b>	
Fair/poor	56 (15.8)
<b>Brand name of the first COVID-vaccine taken</b>	
Coronavac	81 (22.8)
Pfizer	176 (49.6)
Astrazeneca	98 (27.6)
<b>What is your best guess as to whether you will take the second dose of the Coronavirus vaccine?</b>	
Definitively or probably will take	286 (80.6)
Not sure	44 (12.4)
Probably or definitively will not take	25 (7.0)

\*Unless otherwise indicated, data are the respondents' number and percentage (). Percentages may not total to 100 owing to rounding.

Economic considerations suggest that mobile phone text messages are feasible to recruit survey participants. The research was very easy to conduct, manageable by one graduate student as a research assistant committing 14 hours over seven days. This method is also feasible for rapid recruitment of a large sample, from a phone list of 4667 persons, with a mean age of 35.05 years (+/-14.19). We had 355 eligible respondents with a mean age of 33.97 years (+/-14.99). We had some intangible but inexpensive costs with mobile phone devices and web-based services, and the acquisition of 4 prepaid Subscriber Identification Module cards, for about U\$ 10 dollars plus the costs for a private bulk text message service, for U\$ 100 dollars. The total amount was around US 110 American dollars or U\$ 0.31 per respondent. Thus, the method has a considerably lower cost per respondent.

The population and respondents' comparison are reported in Table 2, using the Chi-squared tests. There was a statistically significant difference (*P-value*:.001) between the sample and population distributions for the AstraZeneca and Pfizer groups. All other groups were consistent with the population. Likewise, suggests that they are representative of the population.

**Table 2. The population and respondents group comparison according to Chi-squared goodness of fit calculation.**

Characteristics	Population	%	Respondents	%	Qui square	<i>P-value</i>
	N= 4667		N=355			
Female	1935	41.5	153	43.1		
Male	2732	58.5	202	56.9	.364	.546
18-29 years	2070	44.4	179.0	50.4	1.808	.178
30-44 years	1538	33.0	96	27.0	2.781	.095
45-59 years	760	16.3	53	14.9	.325	.568
≥ 60 years	299	6.4	27	7.6	.679	.409
Morbidities	512	11.0	56	15.8	5.832	.157
Coronavac	908	19.5	81	22.8	1.544	.214
AstraZeneca	2075	44.5	98	27.6	16.835	.001
Pfizer	1684	36.1	176	49.5	10.920	.001

The percentages may not total 100 owing to rounding.  
*P-value* <.05

Table 3 demonstrated Welch's test for comparing population and respondent age groups. The Mean age of the respondent group was 33.97 (SD 14.99), and 35.05 (SD 14.19) of the population, with a *P-value* of .192, and a Cohen's d of 0.0754, corresponding to a small effect size between groups.

**Table 3. Welch's test for the comparison between population and respondent age groups.**

	Respondents N=355	Population N=4667
<i>Mean age</i>	33.97	35.05
<i>Standard Deviation</i>	14.99	14.19
Welch's t-test		
<i>t</i>	-1.3059	
<i>df</i>	403.7489	
<i>P-value</i>	.1923	
Cohen's <i>d</i>	0.0754	
Cohen's $U_3$ (%)	53.0041	
Overlap %	96.9937	
Superiority %	52.1252	
N Needed to Treat	45.9465	

*P-value* <.05

## Discussion

This study investigates whether mobile phone-based advertising can be used to feasibly recruit a representative sample of adults and older adults in a municipality of Brazil to complete a health survey about hesitation for the COVID-19 vaccine. Representativeness was assessed and confirmed by comparing our sample's numerous sociodemographic and health characteristics with the underlying population. Feasibility was confirmed based on an assessment of costs, ease of use, and recruitment time. Moreover, this is an unprecedented study to investigate mobile phone-based text messages for the recruitment of older adults. The vast majority of the national and international literature on social media recruitment focuses on the youth and young adults, some focusing on middle-aged adult populations. (8–11) The reported good participation of the older respondents' results should be of considerable interest to academic researchers, community organizations, and governments because they suggest this method can dramatically reduce research barriers with minimal sacrifices to representativeness.

Although the low response rate, the mobile phone-based recruitment sample was representative of the population for age, gender, and self-reporting comorbidities. Targeting by age and gender was effective in increasing representation. We must assume an over-representation of people with higher levels of education, but targeting advertisements proved both costly and ineffective at improving the representation of participants with lower levels of educational attainment. The response rates of all survey methods decreased over the years, possibly due to a proliferation of web-based questionnaires. (12)

Our reporting contributes to the evidence-based use of online text message surveys for health research. This recruitment method works effectively in the context of the Assis city population. Furthermore, we have shown that online text messages can recruit older adults to research surveys, where the vast majority

of previous research on this subject has only considered younger populations.  
(13)

Caution should be used to consider the first dose vaccine groups from AstraZeneca and Pfizer results. Although significant, there is a potential possibility that these groups have a particular related interest to respond or not to the survey. Therefore, we emphasize the need to proceed with further investigations to elucidate these findings.

The research was very easy to conduct and manageable by one graduate student committing 14 hours per week as a research assistant. This method is also feasible for the rapid recruitment of a large sample, with 355 respondents being recruited in only seven days. A major advantage over other recruitment methods is that it was easy to use targeted advertisements to improve a sample's representativeness and recruit hard-to-reach populations. There was also essentially no footwork involved, such as putting up recruitment posters, and we believe there was less selection bias than using email listservs or snowball sampling. (14,15) Otherwise, there were no risks for the online researchers and the target population in transmitting diseases through face-to-face surveys, especially during the pandemics. (16)

We also found that a mobile phone-based survey is an economical means of recruiting participants for survey research. It has the potential to be a relatively less expensive and timely method of collecting survey information. (17) But, again, maybe due to the higher population interest in the vaccine as an issue that affects everyone.

Limitations of our study were based mostly on subjective interpretations about the feasibility. As our analysis compares representativeness with the population, it is impossible to directly compare the representativeness of samples recruited with mobile phone-based with different recruitment methods. Another perceived limitation was the population's distrust in receiving non-official and unexpected phone messages. Moreover, we can not account for the real number of delivery messages due to limitations as internet failure, missing mobile phone numbers, or interrupted plan services, especially in the low-income population.

We believe there are many potential applications of mobile phone-based advertising for researchers, governments, and community-based organizations who wish to learn about the populations they serve. Besides, we believe that these institutions should previously officially advertise the target population to promote a wider adherence and confidence to respond to the survey.

## **Conclusion**

We achieved a representative sample of respondents using the mobile phone-based survey in Brazil. Furthermore, it was representative in all sociodemographic and health characteristics assessed. Finally, these findings suggest the method is a highly feasible and economical means of recruiting for survey research.

Conflicts of Interest: The authors declare no conflicts.

Fundings: none

## REFERENCES

1. Batterham PJ. Recruitment of mental health survey participants using Internet advertising: content, characteristics, and cost-effectiveness. *Int J Methods Psychiatr Res* [Internet]. 2014 Jun [cited 2022 January 06];23(2):184–91. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/mpr.1421>
2. da Costa TM, Salomão PL, Martha AS, Pisa IT, Sigulem D. The impact of short message service text messages sent as appointment reminders to patients' cell phones at outpatient clinics in São Paulo, Brazil. *Int J Med Inform* [Internet]. 2010 Jan [cited 2022 January 11];79(1):65–70. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1386505609001336>
3. Banskota S, Healy M, Goldberg EM. 15 Smartphone Apps for Older Adults to Use While in Isolation During the COVID-19 Pandemic. *West J Emerg Med*. 2020 Apr;21(3):514–25.
4. Raven MC, Kaplan LM, Rosenberg M, Tieu L, Guzman D, Kushel M. Mobile Phone, Computer, and Internet Use Among Older Homeless Adults: Results from the HOPE HOME Cohort Study. *JMIR mHealth uHealth* [Internet]. 2018 Dec 10 [cited 2022 Jan 11];6(12):e10049. Available from: </pmc/articles/PMC6305882/>
5. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* [Internet]. 2004 Sep 29 [cited 2022 Jan 11];6(3):e34. Available from: </pmc/articles/PMC1550605/>
6. (PDF) Welch's t-test for comparing two independent groups: An Excel calculator (1.0.1) [Internet]. [cited 2022 January 11]. Available from: [https://www.researchgate.net/publication/332217175\\_Welch%27s\\_t-test\\_for\\_comparing\\_two\\_independent\\_groups\\_An\\_Excel\\_calculator\\_101](https://www.researchgate.net/publication/332217175_Welch%27s_t-test_for_comparing_two_independent_groups_An_Excel_calculator_101)
7. Cohen J. Quantitative Methods in Psychology: A Power Primer. *Psychol Bull* [Internet]. 1992 [cited 2022 Jan 11];112(1):155–9. Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.1043.9095>
8. Whitaker C, Stevelink S, Fear N. The Use of Facebook in Recruiting Participants for Health Research Purposes: A Systematic Review. *J Med Internet Res* [Internet]. 2017 Aug 28 [cited 2022 Jan 11];19(8):e290. Available from: </pmc/articles/PMC5594255/>
9. Staffileno BA, Zschunke J, Weber M, Gross LE, Fogg L, Tangney CC. The Feasibility of Using Facebook, Craigslist, and Other Online Strategies to Recruit Young African American Women for a Web-Based Healthy Lifestyle Behavior Change Intervention. *J Cardiovasc Nurs* [Internet]. 2017 Jul [cited 2022 January 11];32(4):365–71. Available from: <https://journals.lww.com/00005082-201707000-00009>
10. Shaver LG, Khawer A, Yi Y, Aubrey-Bassler K, Etchegary H, Roebathan

- B, et al. Using Facebook Advertising to Recruit Representative Samples: Feasibility Assessment of a Cross-Sectional Survey. *J Med Internet Res* [Internet]. 2019 Aug 19 [cited 2022 Jan 6];21(8):e14021. Available from: <https://pubmed.ncbi.nlm.nih.gov/31429409/>
11. Sbaffi L, Rowley J. Trust and Credibility in Web-Based Health Information: A Review and Agenda for Future Research. *J Med Internet Res* [Internet]. 2017 June 19 [cited 2022 January 06];19(6):e218. Available from: <https://www.jmir.org/2017/6/e218>
  12. Sammut DR, Griscti DO, Norman PIJ. Strategies to improve response rates to web surveys: A literature review [Internet]. Vol. 123, *International Journal of Nursing Studies*. Pergamon; 2021 [cited 2022 Jan 12]. p. 104058. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0020748921002054>
  13. Korkeila K, Suominen S, Ahvenainen J, Ojanlatva A, Rautava P, Helenius H, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol* [Internet]. 2001 [cited 2022 Jan 12];17(11):991–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/12380710/>
  14. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. *Comput Human Behav* [Internet]. 2010 Mar 1 [cited 2022 Jan 11];26(2):132–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0747563209001708>
  15. Etikan I. Comparison of Snowball Sampling and Sequential Sampling Technique. *Biometrics Biostat Int J* [Internet]. 2016 Jan 2 [cited 2022 Jan 11];3(1). Available from: <https://medcraveonline.com/BBIJ/BBIJ-03-00055.php>
  16. Newson M, Zhao Y, Zein M El, Sulik J, Dezechache G, Deroy O, et al. Digital contact does not promote wellbeing, but face-to-face contact does: A cross-national survey during the COVID-19 pandemic. *New Media Soc* [Internet]. 2021 Dec 7 [cited 2022 Jan 12];146144482110621. Available from: <https://doi.org/10.1177/14614448211062164>
  17. Vecino-Ortiz AI, Nagarajan M, Katumba KR, Akhter S, Tweheyo R, Gibson DG, et al. A cost study for mobile phone health surveys using interactive voice response for assessing risk factors of non-communicable diseases. *Popul Health Metr* [Internet]. 2021 Dec 28;19(1):32. Available from: <https://doi.org/10.1186/s12963-021-00258-z>