

RESEARCH ARTICLE

The Design and Methodology of the Ohio COVID-19 Survey

Marcus E. Berzofsky¹; Naomi Freedner¹; Caroline Scruggs¹; Robert Ashmead²; Timothy Sahr²; Bo Lu³; Michael Nau²; Eric Seiber⁴; Hilary Rosebrook²; Tom Duffy¹

¹RTI International, Research Triangle Park, NC

²Ohio Colleges of Medicine Government Resource Center, Columbus, OH

³Division of Biostatistics, The Ohio State University, Columbus, OH

⁴Division of Health Service and Management Policy, The Ohio State University, Columbus, OH

Corresponding Author: Robert Ashmead, 150 Pressey Hall, 1070 Carmack Road, Columbus, OH 43210, (614) 685-6817, robert.ashmead@osumc.edu

Submitted December 14, 2020 Accepted February 11, 2021 Published June 21, 2021 <https://doi.org/10.18061/ojph.v4i1.8068>

ABSTRACT

Background: Governments worldwide are balancing contrasting needs to curtail the toll that coronavirus disease 2019 (COVID-19) takes on lives and health care systems and to preserve their economies. To support decisions, data that simultaneously measure the health status of the population and the economic impact of COVID-19 mitigation strategies are needed. In the United States, prior to the onset of COVID-19, surveys or tracking systems usually focused on public health or economic indicators, but not both. However, tracking public health and economic measures together allow policy makers and epidemiologists to understand how policy and program decisions are associated. The Ohio COVID-19 Survey (OCS) attempts to track both measures in Ohio as one of the first statewide population surveys on COVID-19. To achieve this there are several methodological challenges which need to be overcome.

Methods: The OCS utilizes a representative panel offering both cross-sectional and longitudinal analyses. It targets 700 to 1000 respondents per week for a total of 12 600 to 18 000 respondents over an 18-week period. Leveraging a sample of 24 000 adult Ohioans developed from a statewide population health survey conducted in fall 2019, the OCS produces weekly economic and health measures that can be compared to baseline measures obtained before the COVID-19 pandemic began.

Results: The OCS was able to quickly launch and achieve high participation (45.2%) and retention across waves.

Conclusion: The OCS demonstrates how it is possible to leverage an existing health-based survey in Ohio to generate a panel which can be used to quickly track fast-breaking health issues like COVID-19.

Keywords: COVID-19; Panel survey; Survey design; Ohio COVID-19 survey

INTRODUCTION

Understanding the Problem

Governments worldwide are balancing contrasting needs to curtail the toll that coronavirus disease 2019 (COVID-19) takes on lives and health care systems and to preserve their economies. In the United States (US), the federal government has instructed each state to determine the best time to implement its suggested guide-

lines for either the entire state or regions of the state.¹ To make an informed determination, states need detailed regional information about COVID-19 health and economic impacts. However, existing information at the state or county levels is consigned to tracking COVID-19 prevalence (eg, the Johns Hopkins Coronavirus Resources Center) or economic impact data (eg, the US Current Population Survey). Although these individual sources are useful, having health and economic data from a single source would help





states to determine the potential impact of any policy or program on health and economic well-being. Furthermore, because the pandemic is dynamic and changes quickly, data on population risks are outdated in less than a week. To understand the impacts of policy and program decisions, governments must be able to measure change before, during, and after new policies and programs are implemented.

To help inform these decisions and understand how they impact Ohioans, the state of Ohio launched the Ohio COVID-19 Survey (OCS). The OCS leverages a prior Ohio, statewide population survey from which a panel of prospective survey participants was built. The OCS is a weekly web- and telephone-based tracking survey that is representative of Ohio residents and is designed to provide information on health and behavioral measures such as the respondent's COVID-19 testing status and compliance with social distancing and economic measures such as employment status and consumer confidence. This paper explains the OCS methods, which ensure it is generalizable to the Ohio population, and provides paradata from each wave of implementation.

Panel Surveys

Studies measuring emergent public health related topics such as COVID-19 face a challenge in trying to quickly recruit representative respondents who are willing to participate. This is especially true of studies which track health measures over time. To overcome this challenge, survey methodologists can turn to an existing panel to recruit study participants. Panels come in 2 varieties: (1) opt-in or voluntary panel, and (2) probability-based panel.

Opt-in or voluntary panels consist of a collection of persons who chose to be part of a panel and subsequently choose which surveys they participate in. Opt-in panels are not representative of the population for 2 reasons: (1) they skew toward high internet users, and (2) they usually suffer from low (ie, single digit) response rates.² For these reasons, opt-in panels were not considered for the OCS.

Probability-based panels are sets of persons randomly recruited to be a part of a panel. A non-web-based method is used for the recruitment (eg, a random digit dial (RDD) or address-based sample) in order to ensure non-internet users are included. Therefore, these panels are generally representative of the entire population. Probability-based panels require a 2-stage process to conduct a survey. The recruitment stage is first, and the participation stage follows where panel members are invited to take a particular survey. Because of this multi-stage process, response rates are presented for recruitment and participation. The recruitment response rate is the product of the recruitment stage and participation stage. These rates can be low and survey weights are required at each stage to correct for potential biases.^{3,4} However, the participation stage response rates can be high depending on the survey topic.⁵ National probability-panels such as the Understanding America Study (UAS) have been used to quickly pivot and study

emergent topics such as COVID-19.⁵ However, these national panels are relatively small (between 10 000 and 60 000 members spread across all 50 states). Therefore, at the state level, it is difficult for the national panels to obtain a large enough sample to produce reliable estimates. The OCS used a probability-based panel approach, but, to generate reliable estimates, developed its own panel in order to ensure adequate coverage of the entire state.

METHODS

Setting

The OCS is a general population survey of residents of Ohio. The survey was conducted via web and telephone by RTI International.

Design

The OCS has 3 overall analytic objectives: (1) estimate Ohio statewide and regional health and economic indicators related to COVID-19, concentrating on how they change over the progression of the pandemic; (2) understand how individual health and economic statuses and behaviors change over time; and (3) compare current health and economic statuses to prepandemic statuses. To achieve these analytic objectives, the OCS employed a rotating panel design with a 10-minute web-/telephone-based survey.

Panel construction. The OCS sampling frame is a statewide representative panel of Ohioans developed from respondents to the 2019 Ohio Medicaid Assessment Survey (OMAS). The OMAS is a biannual survey of Ohio residents that collects data on health insurance status, health statuses, access to care, determinants of health, and demographics.⁶ The 2019 OMAS, conducted from September to December 2019, obtained 30 068 adult interviews using a dual cell phone and landline RDD frame. The OMAS is weighted to be representative at the statewide and Ohio regional levels. At the end of the OMAS survey, respondents were asked if they were willing to be recontacted. Among the respondents, 24 029 (79.9%) agreed to be recontacted; of these, 16 438 (68.4%) provided telephone and email contact information and 7591 (31.6%) provided only a telephone number. Because not all OMAS respondents agreed to be recontacted, analysts adjusted survey weights for the panel members using a generalized exponential model (GEM) with key demographic and health characteristics to correct for potential panel selection bias.⁷ After this panel inclusion adjustment, the design-based weights for each panel member in OCS fully represent the state and subdomain populations within the state.

Sample design. To achieve the 3 analytic goals, we used a rotating panel design which allowed the OCS to obtain 18 weekly, statewide, cross-sectional estimates and up to 3 repeated interviews with panel members.⁸ Under this design, the full panel was randomly split into 6 rotation groups of approximately 4 000 panel members each. Each rotation group was further randomly split into 20 replicates of 200 panel members each. The weights for each replicate were adjusted to represent the full population by multiplying the design-based weights by 120 (the total number of replicates created). Each week for the first 6 weeks a set of repli-



cates (up to 20) from a new rotation group was released to the field. After the first 6 weeks of data collection the rotation groups were released again. This process was repeated over 3 panel waves. Figure 1 illustrates the rotating panel design for the OCS. The number of replicates released for a weekly rotation group week was the amount anticipated to achieve 700 to 1000 interviews. The design allows for a floor response rate of 17.5% to achieve the minimum collection of 700 weekly interviews. If some replicates were not released, they were held in reserve for future waves to account for possible attrition in the cross-sectional response rates. In waves 2 and 3 all replicates released in wave 1 were released plus any additional replicates needed to achieve the weekly target sample size, estimating for attrition. Therefore, the wave 1 sample release was the set of possible longitudinal respondents that were used for the second analytic objective.

Participants

The OCS target was 700 to 1000 interviews over a 7-day period (weekly) for 18 weeks of data collection split across 3 waves. A rotation group sample was released each Monday.^a The rotation groups did not overlap in the field. An initial invitation was sent by text and email (if available). Three text/email reminders were sent between Tuesday and Wednesday morning. Telephone calling began on Wednesday for all sample members who had not yet responded via web. Up to 3 call attempts per sample member were made through Sunday evening. All panel members who indicated they still lived in the state of Ohio were eligible to take the survey. All respondents were offered a \$5 incentive for completing the survey. Wave 1 was collected from April 20, 2020, until May 31, 2020; wave 2 was collected from June 1, 2020, to July 19, 2020; and wave 3 was collected from July 20, 2020, to August 30, 2020.

Measures/Outcomes

The OCS instrument was a 10-minute survey that obtained critical health, behavioral, and economic indicators. Table 1 details the topics covered in the OCS. To achieve the third analytic objective, the OCS included 5 items from the 2019 OMAS. These items, itali-

cized in Table 1, include health and economic indicators such as self-rated health statuses and food insecurity. In wave 1, the experience of symptoms reference period for the OCS was March 1, 2020. Note that the reference period for time associated questions was modified to “in the past 30 days” for fielding waves 2 and 3.

Statistical Analysis

Two types of analytic weights were created to allow for generalizable inference to the Ohio population: (1) cross-sectional weights, and (2) longitudinal weights.

The cross-sectional weights were produced after each weekly rotation group release. The design-based weights had 3 adjustments made. First, a rotation group and replicate release adjustment was implemented. Because each replicate has weights that represent the full state, the initial replicate weights for each weekly release were adjusted by dividing the weights by the number of replicates released that week. Second, a nonresponse adjustment was made. Because the OCS panel was constructed from the 2019 OMAS, a rich set of characteristics exists and was used to adjust for potential nonresponse biases, including demographic, geographic, financial (income), employment, and health characteristics. Third, a post-stratification adjustment was made for any potential coverage error caused in the creation of the panel rotation groups. Because the design-based weights were not equal (although the expectation was that each randomly created rotation group represents the full population), some variation in the weight totals may exist. Therefore, the weights adjusted for nonresponse were post-stratified to the Ohio population using 2018 American Community Survey estimates.^b The nonresponse adjustment and post-stratification adjustment was conducted using GEM, an iterative raking procedure.

The longitudinal weights were produced after wave 3 was completed among the set of respondents who participated in all 3 waves. The base longitudinal weight was the wave 1 weight for each rotation group because they are the set of sampled persons eligible for the longitudinal analysis. Within each rotation group, a

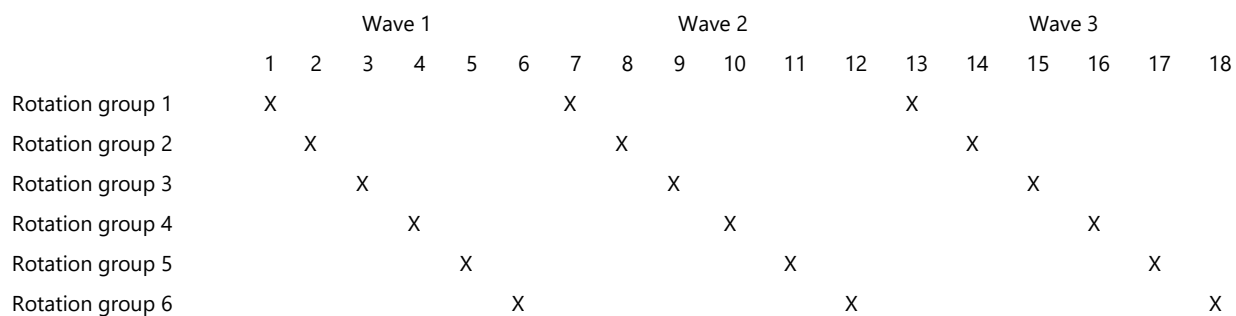


Figure 1. Rotating Panel Design for the OCS by Wave and Week

^aIn wave 2 and wave 3 sampled persons who opt out of the study or refuse to participate in an earlier wave will not be contacted again.

^bThe year 2018 is the most recently available year for American Community Survey data.



Table 1. Topics Covered in the OCS

Topic	Subtopic(s)
Illness	<i>Self-rated health status</i> ; experiencing cough, shortness of breath, chest pain, nausea, vomiting, diarrhea, headache, sore throat, fatigue, difficulty sleeping, unexplained aches and pains, runny or congested nose, loss of smell/taste, dizziness, eye pain, or fever
Mental health	Self-rated interest in doing things; self-rated feeling down or depressed
COVID-19 health-seeking behaviors and testing	Question asks: Because you thought you had COVID-19, have you: had an appointment with health care provider at office or by phone or computer; gone to a testing site or urgent care/ER/hospital; been tested for COVID-19? If tested: date, type, and result of test
Household risk	Anyone in household tested positive for COVID-19
Employment and health insurance	Loss of job because of COVID-19; being paid for work or while staying at home (excluding unemployment insurance); employment situation if employed, including working at home or working away from home, concern about losing job; <i>currently covered by any health insurance</i> ; <i>currently covered by Medicaid</i>
Social behaviors	Are you: staying home; avoiding visiting neighbors, friends, or relatives who don't live with you; avoiding letting neighbors, friends or relatives who don't live with you come to your house; staying 6 feet away from people you don't live with; wearing a face covering when inside a store; wearing a face covering when outside; attending gatherings (other than work) with 10 or more people?
Consumer confidence	Concern for you or your family about: physical health; mental health; personal finances; paying mortgage rent or utilities; <i>worried about running out of food in past 30 days</i> ; <i>run out of food (in past 30 days)</i> ; gotten take-out from a restaurant (since March 15); how you think you and your family will be financially in next 6 months; how COVID-19 has affected you (open ended)

nonresponse and post-stratification adjustment was applied to each set of wave 1 sample members who remained in the study at waves 2 and 3.

In this paper, to evaluate the methods, the statistical analysis utilized paradata. Paradata are data about the process by which the data was collected.⁹ Paradata are useful in assessing the quality and representativeness of the sample. In our case, we evaluated the methods through 5 paradata measures. First, response rates (both the participation response rate and recruitment response rate) were calculated. Second, the attrition rate (the percentage of respondents in a wave that also responded in the subsequent wave) were used to evaluate the power of the longitudinal analysis weights. Third, the distribution of respondents by interview mode (web or telephone) was assessed across waves. Fourth, the refusal rate (the percentage of persons who explicitly declined to take the survey) was used to assess bias in terms of who participated. Fifth, the ineligibility rate (the percentage of the sample who moved out of Ohio since the prior interview wave) was used to determine whether the panel, as time went on, still represented the state of Ohio.

RESULTS

The OCS was launched on April 20, 2020, and completed on August 30, 2020. Across the 3 waves the OCS obtained 17 032 interviews. Table 2 presents the disposition across all 18 weeks and by wave (all rotation groups combined). The overall participation stage response rate was 45.2% yielding a recruitment response rate of 10.0% (ie, =22.2% × 45.2%).¹⁰ By wave, the participation stage response rate was the highest in wave 1 (53.9%) and decreased in

⁶The response rate in the 2019 OMAS was 22.2%.⁶

wave 2 and wave 3 to 42.4% and 39.2%, respectively, due to panel attrition. The wave specific conditional response rates were consistent within each wave's rotation groups. The attrition rate was 66.7% between wave 1 and wave 2 and 75% between wave 2 and wave 3.

The majority of respondents chose to answer via the web mode with an average 58.0% of respondents selecting this mode across the 3 waves. The percentage of web respondents increased in waves 2 and 3 compared to wave 1 (54.1% in wave 1 compared to 60.1% in wave 2 and 61.2% in wave 3).

The refusal rate, the percentage of sample who explicitly indicated they did not want to take the survey, was consistent across the waves, averaging 34.6% and varying by less than 1.2% in any given wave. However, the ineligibility rate increased across waves from 3.0% to 5.5% to 7.1%. This increase in the ineligibility rate is likely due to sample members moving out of Ohio during the 4-month data collection window.

DISCUSSION

The OCS is an example of how prior statewide surveys can be leveraged to develop a panel to track and measure the impact of COVID-19 on health, behavioral, and economic indicators. Three benefits worth highlighting are that: (1) with a panel who recently agreed to participate in a survey and the high saliency of COVID-19 as the topic, the response rate for the study is much higher compared to other statewide surveys; (2) our methodology allows for both a cross-sectional time series and longitudinal analyses; and (3) this survey can be linked to the 2019 OMAS allowing for health and economic comparisons to a prepandemic period estimate.



For rapidly emergent health topics, like COVID-19, where quick collection of data is critical to understanding and tracking how an event impacts the population of interest, having access to an available panel can be very important to the success of a study. If no panel were available and a new sample needed to be recruited the data collection time would be prohibitive to tracking analyses. For example, the 2019 OMAS, which is the source of the panel used in the OCS, took 5 months to recruit 30 000 participants.⁶ Therefore, what made the OCS successful was the inclusion of a question asking OMAS participants if they could be recontacted in the future. Having that agreed upon in advance allowed for the OMAS respondents to be used as a panel, which facilitated OCS data collection within 3 weeks of study inception.

Another key issue in the design is how to best mitigate attrition rates. We utilized 2 methods, reducing the number of recontacts and capping the number of waves. Our study chose to recontact persons every 6 weeks. We believed this time interval would allow us to understand how the pandemic's impact changed for each panel wave without overly burdening the respondents, which can accelerate attrition. Additionally, we capped the panel at 3 waves. While the panel could have been maintained for longer than 4 months, attrition rates would have continued to accelerate making the representativeness of the panel less useful.

As the COVID-19 pandemic continues through the fall of 2020 and into 2021, the OCS has continued with a second panel. Because of the high participation stage response rate, only 13 214 of the 24 000 panel members were invited (Table 2). Therefore, there are approximately 11 000 panel members who could participate in a second OCS iteration without concern of prior panel fatigue.

The main limitation of the design is the low recruitment response rates. While these rates are in line with or slightly higher than national probability-based panels, the rates are still considered low by most survey standards. To mitigate the impact of this limitation we implemented a robust weighting methodology which utilized several correlated outcomes from the OMAS instrument to better calibrate the survey weights. Having health and socioeconomic measures which are typically not available for weighting models, and which are highly correlated to the severity of COVID-19 and its economic impact, can greatly reduce the impact of non-response bias.¹¹

PUBLIC HEALTH IMPLICATIONS

The OCS is providing Ohio the data it needs to understand how COVID-19 is affecting its residents' health and economic welfare. The study continues to produce statewide and sub-state estimates over time, allowing state officials to factor for how policy changes may affect residents in different parts of the state during a pandemic. The OCS maintained a high participation stage response rate (over 45%) across the 3-wave period, demonstrating that a panel can be quickly constructed to conduct a survey on a salient topic like the COVID-19 epidemic to produce estimates that can be immediately operationalized for policy and resource decisions.

REFERENCES

1. White House Coronavirus Task Force. *Guidelines opening up America again*. 2020. Accessed April 26, 2020. <https://www.whitehouse.gov/openingamerica/>
2. Baker R, Blumberg SJ, Brick JM, et al. Research synthesis: AAPOR report on online panels. *Public Opin Q*. 2010;74(4):711-781. <https://doi.org/10.1093/poq/nfq048>

Table 2. Disposition Status and Response Rates for the OCS by Wave

Disposition	Total	Wave 1	Wave 2	Wave 3
Released sample	35 700	13 214	11 676	10 810
Total completes	17 032	6894	5299	4839
Web	9879	3732	3187	2960
Telephone	7153	3162	2112	1879
Breakoffs	807	388	246	173
Web	161	95	41	25
Telephone	646	293	205	148
Nonrespondent (refusal)	12 351	4468	4185	3698
Web	370	189	105	76
Telephone	11 981	4279	4080	3622
Nonrespondent (unknown eligibility)	3699	1069	1299	1331
Ineligible	1811	395	647	769
Participation stage response rate	45.2%	53.9%	42.4%	39.2%
Recruitment response rate	10.0%	12.0%	9.4%	8.7%

Note. The total counts are cumulative across the 3 waves and do not represent unique persons.



3. DiSogra C, Callegaro M. Metrics and design tool for building and evaluating probability-based online panels. *Soc Sci Comput Rev*. 2016;34(1):26–40.
<https://doi.org/10.1177/0894439315573925>
4. Fahimi M. Recruitment-level nonresponse bias analysis for online panels. Paper presented at CIPHER 2020, Washington, DC. Accessed from https://cesr.usc.edu/cipher_2020_agenda
5. Kapteyn A, Angrisani M, Bennett D, et al. Tracking the effects of the COVID-19 pandemic on American households. *Surv Res Methods*. 2020;14(2):179–186.
<https://doi.org/10.18148/srm/2020.v14i2.7737>
6. Berzofsky ME, Duffy T, Johnson K. *2019 Ohio Medicaid assessment survey: methodology report*. 2020. Accessed on December 11, 2020
https://grc.osu.edu/sites/default/files/inline-files/2019_OMAS_MethodsReport_final_combined_21Sept20.pdf
7. Folsom, RE, Singh AC. The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN*. 2000; 598–603. Alexandria, VA: American Statistical Association.
8. Lavrakas, PJ. *Rotating panel design. Encyclopedia of survey research methods*. Sage Publications. 2008.
<https://doi.org/10.4135/9781412963947.n500>
9. Groves RM, Heeringa SG. Responsive design for household surveys: tools for actively controlling survey errors and costs. *J R Stat Soc Ser A Stat Soc*. 2006;169(3):439–457.
<https://doi.org/10.1111/j.1467-985X.2006.00423.x>
10. The American Association for Public Opinion Research. *Standard definitions: Final dispositions of case codes and outcome Rates for Surveys. 9th edition*. AAPOR. 2016
11. Groves R, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opin Q*. 2008;72(2):167–189.
<https://doi.org/10.1093/poq/nfn011>