Follow-Up U.S. General Population COVID-19 Vaccine Uptake Survey: Methodology

Surgo Ventures' final survey sample consisted of 1,670 U.S. adults older than 18 run from March 18 to March 25, 2021. Data was collected using a probability-based household panel ([NORC AmeriSpeak](#)). The survey was conducted online and over the phone and in both English and in Spanish. See Table 1. The sample is representative of the U.S. population and weighted to population benchmarks.

Survey Instrument

Using machine learning algorithms (Classification and Regression Trees), we identified 6 questions based on the key variables from our first NORC survey (more details under Segmentation Methodology) that reliably place a given person into one of the 5 personas -- Enthusiast, Watchful, System Distruster, Cost-Anxious, and Conspiracy Believer -- with 90.5% confidence.

Segmentation methodology

In our first survey, we conducted a segmentation analysis using a k-medoid partitioning around medoids (PAM) clustering algorithm (with a Gower distance metric) to identify clusters of individuals that differed on the following seven variables:

1. **Health insurance status** (whether an individual had health insurance)
2. **Cost barriers to medical care** (whether an individual had delayed medical care in the past year because of cost)
3. **Degree to which an individual agreed the COVID-19 vaccine was unsafe**
4. **Degree of worry about COVID-19**
5. **Early adoption** (whether an individual said they would get COVID vaccine in first three months it is offered)
6. **Conspiratorial belief score** (0-3 score; 1 point each for agree with the following statements: vaccine would insert a tracking chip; COVID-19 is caused by ring of people who manipulate world events; COVID-19 is being exploited by government to control people)
7. **Perception of racial fairness in medical system** (agreement with statement that people of your race are treated fairly in a healthcare setting)

These variables were selected for segmentation based on their relationship to self-reported COVID-19 vaccine likelihood observed in predictive models and their actionability in order to identify population groups and effective interventions. After segments were defined, they were then profiled on COVID-19 vaccine uptake likelihood as well as a variety of demographic and other characteristics. Cluster solutions from 3 to 8 groups were explored. The 5-cluster solution was considered most actionable based on differences between segments in vaccine likelihood barriers and perceptions and is reported here.

Percent of individuals in each cluster represents the population-weighted proportion of respondents in each segment.

Below is the demographic makeup of the sample.

Table 1. Sociodemographics of our sample

Group	Sample %
Total	1670
Gender	
Man	49.2
Woman	50.8
Race/Ethnicity	
Latinx	15.4
White	65.7
Black, African or African American	10.9
Asian, Native Hawaiian, or Pacific Islander	2.6
Others	5.5
Age	
18-34	24.6
35-49	22.5
50-64	25.9
65+	27.1