

GenIC Submitted Under

Cognitive Testing and Pilot Testing for the National Center for Chronic Disease Prevention and Health Promotion

OMB Control Number: 0920-1291 Expiration 5/31/2026

1. **Date:** June 30, 2023
2. **Name, CIO/Program:** Machell Town, NCCDPHP\DPH\PHSB
3. **Title of Study:** Pilot Testing of State-Level Prevalence Estimation Using a YouGov Internet Panel
4. **Study Type:** Pilot Testing
5. **Purpose of Study:**

The purpose of this pilot study is to test the feasibility of using a non-probability internet panel in variously-sized (in terms of population) states to calculate stable health-related population prevalence estimates.

The Behavioral Risk Factor Surveillance System (BRFSS) has been conducting a number of methodological pilots over the past three years. Several of the pilots were conducted to fund data collection methods including internet panel estimation of health-related outcomes. The results indicate that state level estimates may be derived from internet panels in larger states. These estimates could be used to supplement the BRFSS data during years when samples are smaller than anticipated and/or if states need immediate information on prevalence for emerging health outcomes. The pilots also indicated that internet panels in smaller states, where there are fewer panel participants, produce less stable estimates. However, it is not known whether internet panels in states which have relatively stable populations and are in the mid-range in terms of population could produce reliable health-related prevalence estimates. We are interested in comparing the ability of panels to produce reliable estimates over a range of state population sizes. In this project, we will assess a range of states from smallest to largest in population.

All of the previous BRFSS internet panel pilots were conducted using probability-based internet panels. The use of a non-probability internet panel for population prevalence estimates has not been tested.

The information from the YouGov panel survey will be used to:

1. Produce state estimates for all indicators in the final dataset, both raw and weighted to each state's population.
2. Compare estimates from the YouGov panel survey to state estimates from the BRFSS survey.

This pilot study will provide information on the feasibility of collecting data to supplement BRFSS by internet panel, especially for mid-size states. Results may be used to prepare and present methodological research papers at professional conferences or for peer reviewed journals.

6. **Respondent Characteristics:**

This pilot will be conducted in six states. The main mid-size state of interest is Alabama, which is ranked 24th in population among the states and which has a slow (approximately 0.5%) population growth rate. Alabama also has a diverse population in terms of race, education and income. Two larger states will be included: Georgia, with a population approximately double that of Alabama, and California, which is the most populous state. For the YouGov internet panel survey, 500 respondents will be included from each of these three states. Three smaller-population states to be included are North Dakota, New Hampshire, and Wyoming. There will be 250 respondents each from North Dakota and New Hampshire. Wyoming, the least-populated state, will have 200 respondents. All respondents must be U.S. residents, 18 years of age or older.

7. Study Methods:

The questionnaire consists of health questions drawn from the BRFSS core questionnaire; see Attachment 1. Questions will include demographics, chronic conditions, preventive practices and risk behaviors. Additional data items from the contractor's core profile of its respondents (demographic factors, general health status, health insurance, and smoking; see Attachment 2) will also be updated and provided to CDC. Data will be collected via an anonymous online questionnaire hosted by the YouGov with data stored on YouGov servers.

YouGov will conduct data collection during the third quarter of 2023, which will be compared to third quarter state BRFSS data. Unweighted and weighted comparisons will be made to estimates at the state level.

YouGov will provide de-identified datasets to CDC with a codebook describing the location of survey variables in the dataset. YouGov will also analyze data and provide CDC with unweighted and weighted results in the form of tabulations, along with reports summarizing the data findings (weighted and unweighted). Weighting will be based on the American Community Survey. Response rate information will also be provided, and a non-response dataset of demographics on those respondents who started but who did not complete the questionnaire. Details of the panel methodology for recruiting participants and weighting data are in Attachment 4.

8. Recruitment and Incentives:

A commercially available sample from a YouGov internet panel will be used. Incentives provided by the internet panel contractor will be provided. The anticipated incentive for this survey is 1,000 points, equivalent in value to \$1.

9. Personally Identifiable Information (PII):

The internet panel contractor (YouGov) maintains identity and contact information for panel members. That information is stored on a protected server that is separate from the one storing the data from this survey. Survey data are recorded with an encrypted panelist identification number. No personally identifiable information (PII) will be provided to CDC and CDC will not attempt to discover the identity of any participant.

10. Informed Consent/Voluntary Participation:

Prior to beginning the questionnaire, participants will see an introductory screen with language to confirm consent. The proposed language is shown in Attachment 3.

11. Analysis of Data:

We will compare responses to the survey questions in Attachment 1 from the YouGov internet panel to BRFSS results using identical questions on a per-question basis. Results from the YouGov internet panel and BRFSS will be compared by calculating the difference between the estimated percentage or prevalence of each demographic factor, condition, preventive practice or risk behavior between the two

surveys, and then by calculating a confidence interval for each difference. This confidence interval will be calculated as the difference between the estimated percentages plus or minus 1.96 times the standard error of that estimated difference. Statistical hypothesis testing will not be used at this time, because the goal is to determine how comparable the estimates are, and not to determine if they meet a pre-specified level of comparability. Because there is to be no hypothesis testing, no power calculations are needed.

Secondary analyses will include calculating cooperation rates, completion rates, and per-item nonresponse rates. Response rates will be assessed by demographics. We will also explore how the internet panel and BRFSS respondent groups compare with respect to demographics such as age, sex, race/ethnicity, educational level, income, and insurance status.

12. Collection Timeline.

All data collection will be completed within 90 days of OMB approval.

13. Burden Table.

Estimates of Annualized Hour and Cost Burden

Type of Respondent	Number of respondents	Number of responses per respondent	Average burden per response (hours)	Total burden (hours)
Alabama Adults	500	1	15/60	125
Georgia Adults	500	1	15/60	125
California Adults	500	1	15/60	125
New Hampshire Adults	250	1	15/60	62.5
North Dakota Adults	250	1	15/60	62.5
Wyoming Adults	200	1	15/60	50
TOTALS	2200			550

Hourly wage rates are based upon the March 2023 average hourly earnings of \$33.20 for private nonfarm payrolls from the Current Employment Statistics survey conducted by the Bureau of Labor Statistics (available at <https://www.bls.gov/news.release/empsit.t19.htm>).