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 3/31/3023

1. Date: 12/21/2020

Name, CIO/Program: Carol Pierannunzi, NCCDPHP/DPH/PHSB

3. Title of Study: Electronic Health Records and Address- Based Sampling Pilots

4. Study Type: Pilot Testing

5. Purpose of Study:

The purpose of this pilot is to test two methods of data collection that may be used to supplement data from The Behavioral Risk Factor Surveillance System (BRFSS). The purpose of the first project which uses electronic health record (EHR) and survey data from the same respondents will be to assess the quality of data from self-reports and determine the feasibility of merging two datasets together to provide more accurate health information at (sub)state-level jurisdictions. The second project will examine data collection methods other than telephone interviews to collect general health information. The purpose of the second project is to assess the feasibility of the use of an address-based sample (ABS) to administer the BRFSS, which has traditionally relied on a sample of working telephone numbers. These projects will be limited in scope in that they will include relatively smaller numbers of respondents and each will be conducted in only one state.

Although these two projects examine different data collection methods and may provide different information for protocols used with BRFSS data collection and post data collection adjustments, they are both proposed to develop plans to ensure high quality data at the state and substate level.

The EHR project will be conducted in Minnesota in collaboration with the Mayo Clinic and other contractors. The Mayo Clinic maintains a large-scale Electronic Health Records (EHR) system which effectively covers the residents of that state. From a sample of 2,500 randomly selected participants from the primary care providers within the Mayo Clinic EHR, the project anticipates interviewing 900 eligible respondents using information available in health records and which is also collected by self-reports during BRFSS interviews. Interviews will be conducted by phone using contact information from persons in the EHR who sought primary care within the past three years. Details on the EHR project are provided in Attachment 1. We will use a subset of the BRFSS for this project. It will take approximately two years to complete, although data collection will take place over a period of a few months.

The information from the EHR study will be used to:

- 1. Determine the quality of self-reported data and differences in self-reported and electronic health record data on the same/similar question.
- 2. Determine whether and how adjustment weights could be applied to survey data to compensate for errors in self-reporting.
- 3. Evaluate differences among persons within the study site who do/do not have primary care providers and examine the bias inherent in using electronic health information to assess population health.
- 4. Produce methodological paradata reports including (but not limited to) selection bias and weighting to populations.
- 5. Determine whether data collected by surveys of EHR participants can be incorporated into the BRFSS dataset and whether statistical adjustment calculations for mode of response are necessary.

The ABS study will be conducted in the state of Texas. Texas was selected because of the diversity of its population, the large number of counties within the state and the variance among communities in terms of urban/rural population densities. The sample size will be 14,000 addresses in the state of Texas based on a goal of roughly 2,500 completed surveys under conservative response projections.

Respondents will be offered small incentives to complete the survey online. Up to 5 contacts will be made with each address. The first three contacts will include invitation letters and/or postcard reminders. The last two contacts will include mailings of paper and pencil version of the survey which respondents may complete and return. Attachment 2 includes all recruitment materials. Attachment 3 provides the questionnaire to be used.

The information from the ABS study will be used to:

- 1. Assess the feasibility of the use of address-based rather than random digit dialing (RDD) telephone samples to conduct (portions of) the BRFSS.
- Assess whether data from an ABS sample could supplement data from respondents using an RDD sample.
- 3. Review paradata on costs per complete, response rates, timely response and other administrative data to determine efficiencies and data quality.

No data from these studies will be used for population prevalence estimation or rigorous analysis of health data.

Overall results will be used to assess methods for improved data quality and efficiency in the BRFSS data collection process. The practice of testing modes for data collection is standard for all large-scale data collection systems such as the BRFSS and other data collection systems housed within the NCCDPHP.

Respondent Characteristics:

Respondents must be U.S. residents, 18 years of age or older and living in a private residence, within of the states (MN for the Electronic Health Record project and Texas for the ABS project) included the pilot study.

6. Study Methods:

The data will be collected via telephone interview for the EHR project and web-based survey with mailed questionnaire follow-up for the ABS project. Data will then be compared to ongoing surveillance using identical questions.

7. Recruitment and Incentives:

For the Electronic Health Record project, participants will be recruited from persons who have sought primary health care within the past three years through the Mayo Clinic system in the state of Minnesota. Participants will be invited to participate and provided with opportunities for refusal during the screening process. Participation is voluntary. Introductory statements from the BRFSS will be used. No incentives are offered to the respondents in this study.

Participants for the ABS study will be recruited through a random selection of household addresses within the state of Texas. A \$2 incentive will be offered to a split sample of addresses. An additional \$5 incentive will be provided once respondents have completed the online survey. In some cases as much as \$10 may be offered to respondents who complete the online survey after the 5th request. Respondents will be selected from within multiple adult households using the next birthday method. See Attachment 2 for examples of all recruitment materials.

8. Personally Identifiable Information (PII):

General demographic characteristics of respondents will be collected and associated with paradata on questions posed to participants and question responses. RDD sample files will include phone numbers. Respondents will not disclose their names or addresses as part of the question process. ABS sample files will include household addresses. CDC will not retain any PII and will not maintain sample files of phone numbers or addresses. Respondents' phone numbers, addresses and other sample information will be kept in files separate from response files and will not be connected to responses. After completion

of the feasibility tests, sample files will be destroyed. Contactor(s) and CDC will be the only entities with access to the dataset(s).

Information in summary form may be used for presentations on methodology. Results may also be used to prepare and present methodological research papers at professional conferences or for peer reviewed journals. No data from the feasibility test will be used to produce prevalence estimates or analyze public health status.

9. Informed Consent/Voluntary Participation:

During the initial screening for Electronic Health Records project, an interviewer will obtain informed consent. The potential participants will be informed that their telephone number was randomly selected and that participation in the study is completely voluntary. The interviewer will explain the nature of the study and approximately how long the survey will take. The potential participant will be told that they do not have to answer any question that they do not want to, and they can stop the survey at any time. They will be informed that their response on the survey will not be connected to any personal information, and it they have any questions they can call the survey point of contact. Since commercially available internet panels are comprised of voluntary participants, there will not be a need to obtain additional informed consent from those potential participants.

The ABS recruitment postcards and letters will inform respondents of the nature of the study. Introductory screens will also inform participants that their participation is voluntary and that they are not required to answer any questions.

No PII will be collected. All telephone numbers and addresses needed for sampling will be retained by contractors for the use of the feasibility experiments and then files will be destroyed. The PII section details how information will be secured.

10. Analysis of Data. Explain how data will be analyzed.

Analyses will include calculating cooperation rates, completion rates, overall response rates, and the distribution of demographic characteristics for persons completing the surveys for the two projects. Data quality will also be assessed, including item nonresponse, miscoded/out of range responses and problems with household-level selection.

Responses from Electronic Health Records will be compared to responses on the same/similar questions from the survey. Comparisons of responses will also be made to ongoing data collection from the BRFSS in the two states. These comparisons will be used to determine whether there are differences between current methods of data collection and methods which may be adopted in the future and/or be used to supplement current methods.

11. Collection Timeline.

The timelines below for the two projects (listing milestones) is in months.

Address Based Sampling Project												
Item	1	2	3	4	5	6	7	8	9	10	11	12
Web-based Programming												
Sampling												
Data Collection and Monitoring												
Data Cleaning												
Analysis												

Electronic Health Records Project												
Item	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24
Questionnaire												
development												
CATI Programming												
Interviewer training												
Data Collection												
and Monitoring												
Data Cleaning												
	1				1							

12. Burden Table.

Analysis Report Writing

Report Writing

Estimates of Annualized Hour and Cost Burden

Type of Respondents	Stage of Survey Administration	Number of respondents	Number of responses per respondent	Average burden per response (in hours)	Total burden in hours	
	Online survey response	2,000	1	.5	1000	
General Texas Adult Population	Paper and pencil survey response	500	1	.5	250	
	Total ABS	2,500			1250	
Minnesota adult residents enrolled in the	CATI Screening for RDD Survey	2,500	1	.1	250	
Mayo EHR who received primary care in	Telephone Interview	900	1	.5	450	
past 3 years	Total EHR				700	
Total					1950	

Table 12B Estimated Annual Cost Burden							
Stage of Survey Administration	Single Administration Burden Hours	Average Hourly Rate*	Total Single Administration Cost Burden				
ABS Online survey response	1000	\$25.72	\$25,720				
ABS Paper and pencil survey response	250	\$25.72	\$6,430				
CATI Screening for EHR respondents	250	\$25.72	\$6,430				
Surveying HER respondents	450	\$25.72	\$11,574				
Total			\$50,154				

^{*}Based upon the average hourly earnings from the Bureau of Labor Statistics May 2019 National Occupational Employment and Wage Estimates (available at https://www.bls.gov/oes/current/oes_nat.htm).