



## **Dental Access Survey**

The Pew Charitable Trusts and Harvard School of Dental Medicine  
September 2018

Method Summary Report prepared by  
David Dutwin, Clay Daniel and Laura Pipes

**Contents**

Contents ..... 2

Sample Design ..... 3

Questionnaire Design and Programming ..... 4

Survey Field ..... 5

Weighting Procedures ..... 6

Margin of Sampling Error ..... 10

Response Rate ..... 11

## Overview

The Pew Charitable Trusts and Harvard School of Dental Medicine contacted SSRS to conduct a survey among respondents who have dental insurance through a private insurer or Medicaid in specific states (Alaska, Arkansas, California, Colorado, Connecticut, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Massachusetts, Michigan, Minnesota, Missouri, Montana, Nebraska, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Vermont, Virginia, Washington, Washington D.C., Wisconsin and Wyoming). These states were selected because as of 2018, their state Medicaid program offers dental coverage to adults that covers more than emergency dental services. Specifically, The Pew Charitable Trusts and Harvard School of Dental Medicine wanted to understand barriers in the dental care delivery system as well as usage of it and attitudes toward future trends. In July of 2018, Pew contracted with SSRS to conduct this survey research.

SSRS completed interviews with 405 respondents with dental insurance through a private insurer and 423 respondents with dental insurance through Medicaid for a total of 828 respondents in specified states. The average length of interview was 13.73 minutes. This report details the methodological components of this study, including the sample design, questionnaire design and programming, field operations, weighting, response rate and data processing and deliverables.

## Sample Design

The study was designed for the majority of interviews to be completed through the SSRS Omnibus, a national, weekly, dual-frame bilingual telephone survey design. This component resulted in 88.5% of the targeted, total number of completes. A custom component, using prescreened Medicaid sample, was added as a supplement to gather the remaining completes—approximately 11.5% of the total—to ensure the study finished in the time frame allowed.

The phone numbers utilized for this study were randomly generated for the OMNI portion and prescreened from the OMNI for the custom Medicaid portion:

- RDD landline sample
- RDD cell sample, and
- Prescreened OMNI Medicaid sample pulled from January to May 2018—1,000 randomly selected respondents in specified states to re-contact for this study

The RDD and Listed landline sample were generated through Marketing Systems Group’s (MSG) GENESYS sampling system. MSG is one of the survey research industry’s largest statistical sampling companies, and a supplier to social science researchers and governmental organizations such as the U.S. Census Bureau and Centers for Disease Control and Prevention. The standard GENESYS RDD methodology produces a strict single stage, Equal Probability Selection Method (epsem) sample of residential telephone numbers. In other words, a GENESYS RDD sample ensures an equal and known probability of selection for every residential telephone number in the sample frame. The sample was generated shortly before the beginning of data collection to provide the most up-to-date sample possible, maximizing the number of valid telephone extensions. Following generation, the RDD sample was prepared using MSG’s proprietary GENESYS Elevate procedure, which identifies and eliminates a large percentage of all non-working and business numbers.

Using a procedure similar to that used for the landline sample, MSG generated a list of cellphone telephone numbers in a random fashion. Inactive numbers were flagged and removed utilizing MSG’s CellWins procedure.

Table One below summarizes the number of interviews completed by quota and sampling frame.

<b>Table One: Distribution of Interviews</b>			
	<b>Medicaid</b>	<b>Private Insurance</b>	<b>Total</b>
<b>Custom Landline</b>	35	--	35
<b>Custom Cell</b>	60	--	60
<b>OMNI Landline</b>	121	156	277
<b>OMNI Cell</b>	207	249	456
<b>Total</b>	423	405	828

**Questionnaire Design and Programming**

The questionnaire was developed by The Pew Charitable Trusts and Harvard School of Dental Medicine in collaboration with the SSRS project team. SSRS provided feedback regarding question wording, order, clarity, and other issues related to questionnaire quality and formatted the survey for programming. SSRS programmed the study into CfMc 8.6 Computer Assisted Telephone Interviewing (CATI) software. Extensive checking of the CATI program was

conducted to ensure that skip patterns followed the design of the questionnaire and that the interview flowed smoothly. Further, thousands of cases of random data were also run through the program to verify program logic accuracy. Once the survey was finalized, the instrument was translated into Spanish, so respondents could choose to be interviewed in English or Spanish, or switch between the languages according to their comfort level. A total of 24 surveys were conducted in Spanish.

## Survey Field

### ***Pretest***

Twenty pretest interviews were completed prior to the field period. The live pretest of the survey instrument was conducted on July 25 and July 26, 2018. SSRS provided recordings and a detailed summary of pretest findings, which included feedback from the interviewers. The final draft of the questionnaire was revised on the basis of the pretest and follow-up tests. Changes were made in order to enhance respondent comprehension of questions and improve screening.

### ***Main Study***

The field period for this study was August 1 through September 9, 2018. All interviews were completed through the CATI system. The CATI system ensured that questions followed logical skip patterns and that complete dispositions of all call attempts were recorded.

CATI interviewers received both written materials on the survey and formal training prior to the study launch. The written materials were provided prior to the beginning of the field period. They included an annotated questionnaire that contained information about the goals of the study as well as detailed explanations with obstacles to be overcome in getting the best possible answers to questions, and respondent problems that could be anticipated ahead of time as well as strategies for addressing the potential problems.

Interviewer training was conducted immediately prior to launch. Call center supervisors and interviewers were walked through each question in the questionnaire. Interviewers were given instructions to help them maximize response rates and ensure accurate data collection. In order to maximize survey response, SSRS also enacted the following procedures:

- At least six call attempts were made to contact non-responsive numbers (no answer, busy, answering machine) before considering a sampling unit “dead”.
- Each non-responsive number was contacted multiple times, varying the times of day, and the days of the week that callbacks were placed using a programmed differential call rule.
- Respondents were offered the option to set a schedule for a callback.
- Interviewers called sample with a suspended disposition in an attempt to persuade respondents to finish the survey.

### **Weighting Procedures**

Weighting is generally used in survey analysis to compensate for sample designs and patterns of non-response that might bias results. For the SSRS Omnibus sample, the original SSRS Omnibus weight was carried over in a self-weight procedure. For the RDD sample, benchmarks were extracted via the SSRS Omnibus portion and weighted to match parameters of adults in the targeted states with Medicaid or private insurance, and respective eligibility criteria.

### ***SSRS OMNIBUS SAMPLE***

Each SSRS Omnibus insert is weighted to provide nationally representative and projectable estimates of the adult population 18 years of age and older. The weighting process takes into account the disproportionate probabilities of household and respondent selection due to the number of separate telephone landlines and cellphones answered by respondents and their households, as well as the probability associated with the random selection of an individual household member. Following application of the above weights, the sample is post-stratified and balanced by key demographics such as age, race, sex, region, and education. The sample is also weighted to reflect the distribution of phone usage in the general population, meaning the proportion of those who are cellphone only, landline only, and mixed users. Specific steps are provided below:

#### **Base weight**

1. The SSRS Omnibus base weight adjustment for probability of selection is based on:

- a. A phone number's probability of being sampled from the landline ( $PN_{LL}$ ) or cellphone ( $PN_{CELL}$ ) sampling frame. This is calculated as:

$$PN_{LL} = S_{LL}/F_{LL}$$

$$PN_{CELL} = S_{CELL}/F_{CELL}$$

Where:

$S_{LL}$  = the size of the landline sample

$F_{LL}$  = the size of the landline sample frame

$S_{CELL}$  = the size of the cell sample

$F_{CELL}$  = the size of the cell sample frame

- b. The probability that a respondent will be selected based on the number of adults in their household. This adjustment is equal to the number of adults in the respondent's household (AD), capped at 3.
- c. Each respondent's access to a landline (LL) and/or cellphone (CP). Respondents reachable by landline were assigned LL=1, those who could not be reached by landline were assigned LL=0. Similarly, respondents who had a cellphone were assigned CP=1 and those without a cellphone were assigned CP=0.

A respondent's probability (PR) of being reached by each phone type was calculated as:

$$\text{Landline: } PR_{LL} = LL \times PN_{LL} / AD$$

$$\text{cellphone: } PR_{CELL} = CP \times PN_{CELL}$$

And the overall probability of selection from either was calculated as:

$$P_{(LL \cup CELL)} = PR_{LL} + PR_{CELL} - (PR_{LL} \times PR_{CELL})$$

The base weight adjustment (BW) for probability of selection was calculated as:

$$BW = 1/P_{(LL \cup CELL)}$$

## Post-stratification

1. With the base weight applied, the sample was weighted to reflect the distribution of known adult-population parameters based on the most recent March Supplement of the U.S. Census Bureau's Current Population Survey (CPS)<sup>1</sup>. Weighting was accomplished using SPSSINC RAKE, an SPSS extension module that simultaneously balances the distributions of all variables using the GENLOG procedure.

<sup>1</sup> Sarah Flood, Miriam King, Steven Ruggles, and J. Robert Warren. *Integrated Public Use Microdata Series, Current Population Survey: Version 5.0*. [dataset]. Minneapolis: University of Minnesota, 2017. <https://doi.org/10.18128/D030.V5.0>.

2. The population parameters used for post-stratification are: age (18-29, 30-49, 50-64, 65+) by gender; Census region (Northeast, North-Central, South, West) by gender; education (less than high school, high school graduate, some college, four-year college or more); race/ethnicity (white non-Hispanic, Black non-Hispanic, Hispanic and born in the U.S., Hispanic and born outside of the U.S.<sup>2</sup>, other non-Hispanic); marital status (married/not married), population density (divided into quintiles) and phone usage (cellphone only, landline only, both).

To handle missing data among some of the demographic variables we employed a technique called hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. These are further determined by variables predictive of non-response that are present in the entire file. We used an SPSS macro detailed in “Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data” (Myers, 2011).

3. Weight truncation (“trimming”): To ensure the consistency of the population estimates produced week-to-week by Excel, the weights will undergo truncation (or “trimming”) so that they do not exceed 4.0 or fall below under 0.25.

The sum of weights will equal the sample N.

### ***RDD SAMPLE (Medicaid under age 65)***

#### **Base weight**

1. Probability of Selection (Pwt): Step 1 as outlined at base-weight section for the SSRS Omnibus sample was undertaken focusing on frame probabilities for the targeted states only.
2. Inverse Probability Weighting (IPW): The callback sample is the product of random digit sampling (for the SSRS Omnibus poll). However, it is susceptible to nonresponse bias stemming from systematic differences between respondents willing to complete a second interview, and those who are not. This is a typical problem for callback studies, where bias is introduced if respondents in the original study are systematically different than those responding to the follow-up survey. Inverse Probability Weighting (IPW) or Propensity Weighting is often used to adjust for attrition in panel/longitudinal studies. Characteristics of the respondents as measured in the initial studies were used to model their probability of response to the callback survey. Propensity was modeled through logistic regression in which the outcome is whether or not they completed the callback screener. The predictive values are calculated as the probability of a person completing the screener. The propensity weight (propwt) was calculated as the inverse of the predicted probability of completing the callback interview calculated by the logistic regression model. Variables typically used in this model are demographics (gender, home ownership, marital status [married, or not], employment status [full-time, unemployed, all other statuses], race [Black, Hispanic, all else], age [younger than 30, 50 or older], educational attainment [high school or less, college or more], income

---

<sup>2</sup> Since this is meant to address the percent of Spanish speakers in the weighted sample, respondents born in Puerto Rico are included with those born outside of the U.S.

[\$25,000 or less, \$75,000 or more, Refused], and population density in county of residence) and behavioral items such as voter registration, and landline usage.

3. Base Weight (Bwt): The base weight will be calculated as the product of the first two steps:

Pwt/IPW

### Post-stratification

1. With the base weight applied, the sample was weighted to reflect the distribution parameters based on the weighted Medicaid under age 65 results from the Omnibus Survey. Weighting was accomplished using SPSSINC RAKE, an SPSS extension module that simultaneously balances the distributions of all variables using the GENLOG procedure.
2. The population parameters used for post-stratification are age (18-29, 30-49, 50-64, 65+); gender; Census region (Northeast, North-Central, South, West); education (less than high school, high school graduate, some college, four-year college or more); and race/ethnicity (white non-Hispanic, Black non-Hispanic, Hispanic, other non-Hispanic).

To handle missing data among some of the demographic variables we employed a technique called hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. These are further determined by variables predictive of non-response that are present in the entire file. We used an SPSS macro detailed in “Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data” (Myers, 2011).

3. Weight truncation (“trimming”): Weights were trimmed to prevent individual interviews from having too much influence on the final results. The sample was truncated at the 96<sup>th</sup> and 4<sup>th</sup> percentiles.

Given the introduction of additional age under 65 Medicaid interviews via the custom RDD re-contact sample, the proportion of age under 65 Medicaid respondents was balanced back to the original distribution in the weighted SSRS Omnibus survey.

Lastly, one final adjustment was made to balance the Medicaid and private insurance groups in proportion with the population.

The sum of weights will equal the sample N.

### Effects of Sample Design on Statistical Inference

Post-data collection statistical adjustments require analysis procedures that reflect departures from simple random sampling. SSRS calculates the effects of these design features so that an

appropriate adjustment can be incorporated into tests of statistical significance when using these data. The so-called “design effect” or *deff* represents the loss in statistical efficiency that results from systematic non-response. The total sample design effect for this survey is 1.50.

SSRS calculates the composite design effect for a sample of size *n*, with each case having a weight, *w<sub>i</sub>* as:

$$deff = \frac{n \sum_{i=1}^n w_i^2}{\left( \sum_{i=1}^n w_i \right)^2} \quad \text{formula 1}$$

In a wide range of situations, the adjusted standard error of a statistic should be calculated by multiplying the usual formula by the square root of the design effect ( $\sqrt{deff}$ ). Thus, the formula for computing the 95% confidence interval around a percentage is:

$$\hat{p} \pm \left( \sqrt{deff} \times 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right) \quad \text{formula 2}$$

where  $\hat{p}$  is the sample estimate and *n* is the unweighted number of sample cases in the group being considered.

The survey’s margin of error is the largest 95% confidence interval for any estimated proportion based on the total sample— the one around 50%. The margin of error for the entire sample is  $\pm 5.0$  percentage points. This means that in 95 out every 100 samples drawn using the same methodology, estimated proportions based on the entire sample will be no more than  $\pm 5.0$  percentage points away from their true values in the population. It is important to remember that sampling fluctuations are only one possible source of error in a survey estimate. Other sources—such as respondent selection bias, questionnaire wording, and reporting inaccuracy— may contribute additional error of greater or lesser magnitude.

## Design Effects and Margins of Sampling Errors

	Number of Interviews	Margin of Error with Design Effect	Design Effect
Total	828	+/- 5.0 percentage points	2.12
Medicaid	423	+/- 5.8 percentage points	1.49
Private Insurance	405	+/- 6.0 percentage points	1.51

## Response Rate

The response rate for this study was calculated using AAPOR's RR3. The Dental Access Survey achieved a response rate of 21.5% on landline and 16.4% on cell for the custom prescreened sample, for an overall re-contact response rate of 18.3%. The response rate for the OMNI sample achieved 5.5% on landline and 4.3% on cell for an overall response rate of 4.7%. Table Two outlines the Dental Access Survey sample disposition:

<b>Table Two: Sample Disposition</b>				
	<b>Custom</b>		<b>OMNI</b>	
<b>Eligible, Interview</b>	<b>LL</b>	<b>Cell</b>	<b>LL</b>	<b>Cell</b>
Complete	35	60	277	456
<b>Unknown eligibility non-interview</b>				
Always busy	3	4	7,325	4,559
No answer	92	158	184,189	164,608
Answering machine-don't know if household	46	107	67,235	155,762
Call blocking	2	5	4,412	43,021
Housing unit, unknown if eligible respondent	99	235	56,523	100,260
No screener completed	10	23	7,722	4,773
<b>Not eligible</b>				
Fax/data line	0	0	19,226	770
Non-working number	24	55	967,804	381,281
Business, government office, other organization	0	2	36,279	4,848
No eligible respondent	29	37	3,027	5,039