

Supporting Statement: Part B

Survey of Primary Care Physicians on Oral Health

Submitted by:

Judith Labiner-Wolfe, PhD

Office on Women's Health

Office of the Assistant Secretary for Health

Department of Health and Human Services

Email: Judith.Labiner-Wolfe@hhs.gov

Phone: (202) 260-0904

May 15, 2012

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Respondent Universe and Sampling Methods
2. Procedures for the Collection of Information
3. Methods to Maximize Response Rates and Deal with Nonresponse
4. Test of Procedures or Methods to be Undertaken
5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

LIST OF ATTACHMENTS

1. Authorizing Legislation: Public Health Service Act
2. Relevant Oral Health Surveys Table
3. Literature Review Reference List
4. Federal Register Notice
5. Contractor's Nondisclosure Agreement
6. IRB Approval
7. Telephone Screener
8. Physician's Oral Health Survey
9. Survey Cover Letter
10. Physician Frequently Asked Questions Letter

PART B.

COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Respondent Universe and Sampling Methods

Target Population

The target population for this survey is primary care physicians whose specialty is either Family Practice (FP) or Internal Medicine (IM), by far the two largest primary care specialties caring for adults. There are several other eligibility criteria for the physicians within these specialties. Physicians in these specialties must see patients as part of their work day lives (i.e., they are not engaged solely in teaching, administration, or research), though they may not be exclusively engaged in patient care activities. Eligible physicians may not be federally employed. Osteopaths are eligible for the survey.

Sample Frame

There are two files that serve as prime candidates for the source of the sample frame of physicians in the specialties of Family Practice (also known as Family Medicine) and Internal Medicine: the American Medical Association (AMA) Master file of physicians and the data bases maintained within the National Plan and Provider Enumeration System (NPPES) that contain the National Provider Identifier (NPI) records. There are two “NPI” data bases, one at the individual level and one at the organization level. We propose to use a subset of the individual level file, and, for ease of reference here, we will refer to it as the NPI file.

Both files are expected to provide a high degree of coverage of the two specialties of interest to this study. The AMA file is based on records established for students entering accredited medical schools (with provision made for identifying physicians practicing in the U.S. but who obtained degrees from foreign medical schools). The NPI file consists of those physicians who have obtained an NPI ID. This file is maintained by the Centers for Medicare and Medicaid Services (CMS). All Health Care Providers who conduct HIPAA standard transactions (e.g., claims and eligibility inquiries) are required to obtain an NPI. By definition a HIPAA standard transaction is electronic (rather than paper, fax, or phone). Electronic claims and electronic eligibility inquiries are examples of such transactions. Thus, the vast majority of Health Care organizations (e.g., hospitals, practices, etc.) are required to do so. All physicians are eligible to obtain an NPI, but unless they themselves conduct HIPAA standard transactions, under the provisions of HIPAA they are not required to obtain an NPI. That said, examination of counts of physicians by specialty indicate (through comparison with AMA totals) that most physicians appear to be obtaining NPI numbers. In addition, the Affordable Care Act (ACA) passed by Congress on March 23, 2010 requires all Medicare and Medicaid providers to obtain NPIs if they are eligible for NPIs. This would be expected to further increase the number of physicians who would obtain NPIs. The table shows counts of those identified as FPs on the NPI file of individuals in June

2010 and May 2011 (around 100,000 in 2010, increasing to about 105,000 in 2011) compared to the same counts for the AMA in September 2010 and January 2012 (just under 100,000 in both years). For IMs the corresponding counts are also close (NPI: about 110,000, increasing to 115,000; AMA: about 120,000, dropping to about 118,500).

Physician Specialty	NPI File Count		AMA Master File Count	
	June 2010	May 2011	September 2010	January 2012
Family Medicine	100,959	105,219	98,441	99,970
Internal Medicine	111,051	115,617	120,157	118,519

Advantages and Disadvantages of the AMA and NPI Files

AMA File Advantages

1. Expected high coverage of the target populations for the study
2. Relatively large number of auxiliary variables available for nonresponse adjustments

AMA File Disadvantages

1. Built in nonresponse—Primary Care Physicians who do not want to be contacted represent perhaps 1-3 percent of a sample, depending on specialty (based on studies Westat has done using the AMA file as a sample frame)
2. Contact information often out-of-date—tracing efforts can be costly (in terms of money, time, and nonresponse)—the NPI data base can help reduce costs when the AMA data proves to be out of date by serving as another source of contact information since it is available on line and is free to use
3. High ineligibility rate due to out-of-date information on employment status
4. Specialty indicators can be out-of-date, resulting in potential bias

NPI File Advantages

1. Expected high coverage
2. Contact information currently more up-to-date than AMA
3. Free (although a data base needs to be established and edited)

NPI File Disadvantages

1. Limited variables for the purposes of nonresponse adjustments and sample stratification (geocoding addresses on the frame and extracting data from other sources can help augment such information)
2. The accuracy of the specialty (taxonomy) is uncertain—CMS does not attempt to verify its accuracy—to the extent that it is inaccurate, there are potential issues of bias

Westat has had substantial experience using the AMA master file for sampling primary care physicians. It has been problematic to use, with out-of-date contact and other information that contribute to concerns about nonresponse and coverage bias. At the Provider Survey Methods Workshop sponsored by the National Cancer Institute in November 2010, David Woodwell of the National Center for Health Statistics (NCHS), a statistician who works on the National Ambulatory Medical Care Survey (NAMCS), reported similar concerns, including misclassification of physician specialties. One slide he presented showed that for the years 2005-2009 of the physicians sampled for the NAMCS over 25 percent of them proved to be out-of-scope. The NAMCS has eligibility criteria similar to this study on Oral Health.

Taking all of this into consideration, we are proposing to use the NPI file as the basis for the sample frame for the HCS. Comparing the numbers of FPs and IMs on the AMA and NPI files, we expect high coverage of the target populations. Moreover, we expect to obtain higher response rates than we would obtain using the AMA file. The NPI variables available for nonresponse adjustment purposes may be somewhat more limited than would be available using the AMA file. The NPI specialties have not been confirmed by CMS. NAMCS has experienced misclassification of specialties, so it is not necessarily the case that the AMA data base represents an improvement in this regard. Special runs done by NCHS for Westat based on NAMCS data for one year have shown that a small percentage (about 4 percent) of responding IMs were sampled from non-IM strata and even a smaller percentage (about 2 percent) of FPs were misclassified in this fashion. The stability of these estimates is uncertain. However, there does not appear to be evidence of substantial concerns about bias resulting from misclassification on the AMA file. It is expected that the NPI file would be more up-to-date and thus perhaps less susceptible to such misclassification, but that is not known.

Population Estimates and the Sample Frame

For this study we plan to use all FP and IM records appearing on the NPI file. In surveys of physicians recently carried out for the National Cancer Institute (NCI) by Westat, Westat has used the NAMCS model, relying on specific variables found on the AMA file to eliminate some physicians from the sample frame. In particular, physicians can be characterized as Office Based, or working in Teaching, Administration, or Research among other classifications. Being classified as Office Based on the AMA file is an eligibility requirement for NAMCS. However, there have been studies indicating that relying on the AMA classification of Office Based may result in undercoverage of a physician specialty since AMA data can be out-of-date and physicians not flagged as Office Based may now be. NAMCS estimates for the number of physicians associated with many specialties are now routinely produced. We plan to compare our estimates of physicians to those obtained from the NAMCS as well as those obtained from the NCI surveys carried out by Westat to assess the degree of undercoverage that may be incurred by excluding Office Based physicians from AMA sample frames (using the NPI file as the source of the frame will confound this assessment to some degree as will the specific definition of those eligible for the Oral Health study, which will differ to some degree to that used for the NAMCS. However, as mentioned before, both the full AMA and NPI files are expected to provide very high coverage of the IM and FP populations, so a comparison of the estimates is expected to be informative.)

Response and Eligibility Rates

As mentioned earlier, Westat has carried out several studies of primary care physicians for the National Cancer Institute (NCI). The overall response rates have varied somewhat depending on the subject matter and specialty, but have been basically in the 60-70 percent range. We expect them to be in the same range for this study as well. We have found that IMs have higher ineligibility rates than FPs. For example, for the Cancer Screening Study (CSS) carried out by Westat for NCI in the fall of 2006, the eligibility rate for FPs was about 83 percent while that for IMs was about 68 percent. For the CSS the sample frame was based on the AMA masterfile and excluded physicians not flagged as office-based on that file. We expect that including all physicians from the NPI file will result in a reduced eligibility rate though the degree of ineligibility is uncertain. As a result, we plan to select a reserve sample, in order to supplement the main sample should initial assumptions depart somewhat from what is experienced in the field

2. Procedures for the Collection of Information

Statistical Method for Stratification

We plan to use several variables from the NPI file (the planned source of the sample frame) for implicit stratification (i.e., sorting) purposes. These include: Taxonomy (specialty); sole proprietorship (yes, no, or unanswered); and geography (derived from address). Sole

proprietorship practices are a subset of solo practices and indicate a particular legal status for the practice. We will use the state indicated in the address to assign a physician to a particular Census region and the addresses will be geocoded in order to assign an MSA status to the physician as well. Gender of provider appears on the NPI data base and will be used if the degree of missing data is not extensive. We plan to use specialty for implicit stratification purposes because analyses are mainly focused on primary care physicians rather than individual specialties. However, we plan to use it as the initial sort variable, ensuring a proportional distribution of the sample across the two specialties with minimum variability in sample sizes for each specialty. This provides a reasonable framework for developing separate estimates for each specialty as well as comparison between specialties, if desired.

Another factor to consider is that the number of IMs on the NPI file is expected to be roughly 10,000 more than the corresponding number of FPs. Thus, sampling from a frame consisting of both FPs and IMs will produce a slightly larger sample of IMs. However, if the product of response and eligibility rates for IMs is somewhat lower than for FPs (as might be expected based on data available from the NAMCS as well as the NCI studies undertaken by Westat), the actual sample yield of eligible respondents for the two could be roughly the same for the two specialties or perhaps even lower for IMs compared to FPs. Since the focus of the study is estimates for primary care physicians who are either FP or IM, this is not a major concern, and we do not plan to vary response rates by specialty to increase the yield for a particular specialty for comparison purposes. Doing so would result in a design effect due to differential sampling rates, serving to reduce precision for the main target population of interest to the study.

Statistical Method for Sample Selection

Our planned sort order to achieve implicit stratification of the physicians on the frame is: specialty; within specialty by region; within region by MSA status; within MSA status by sole proprietorship status; and within sole proprietorship status by the census block group of the physician's contact address. This final sort variable will help reduce the likelihood that physicians in the same practice will be selected after sorting the sample frame, as an equal probability systematic random sample of physicians will be selected. The physician contact information provided on the sample frame will be used to contact the sampled physician.

Sample Size

We are targeting about 600 completed interviews in total. Assuming a screener response rate of .9, a completed interview response rate of .7 among those physicians designated as eligible to receive the main interview questionnaire, and an overall eligibility rate across the screener and main interview of .75, we plan to select 1,300 physicians from the sample frame consisting of FPs and IMs.

Data Collection Procedures

Tracing

It is anticipated that the accuracy of the NPI file will result in a low need for tracing. However, if the sampled physician no longer works at the practice contacted, s/he will be traced until s/he is found. This tracing will take place at any point needed during the field period.

Screener Calls

Based on our previous experience with the AMA Master File, we expect that perhaps up to 50 percent of our initial sample will be ineligible with eligibility rates varying by specialty. Experienced Westat telephone interviewers will conduct screening telephone calls to make an initial determination as to the eligibility of a physician to participate in the study as well as to confirm the office mailing address. Because interviewers will most likely be speaking with an office staff person rather than the physician, the screener questions will cover only the most basic eligibility criteria to eliminate physicians who are deceased, retired, or no longer in care delivery. Physicians who are identified as ineligible will be excluded from further study activities.

Initial Mailings (see attachment 8)

Once a physician is identified as eligible during the screening call, we will send him/her a survey package by Priority Mail. Using the US Postal Service's Priority Mail envelope will set this survey apart from other mailings, making it more likely that the office staff will bring the survey to the physician's attention before it gets lost in the office shuffle. The survey package will contain a personalized cover letter on OWH letterhead (**see Attachment 9**), an information sheet (**see Attachment 10**), the study questionnaire, a postage-paid return envelope and a \$50 check as reimbursement for their time. Offices that have not returned the questionnaires within 13 days after the first mailing will be sent a second survey package by Priority Mail. The second package will contain a more urgently worded cover letter as well as the other materials in the first mailing with the exception of the incentive.

Follow-up Calls

Experienced Westat interviewers will make a follow-up call to non-responding physician offices in order to verify that the package has been delivered to the physician's desk, leave a reminder message for the physician about the importance of the study and provide a toll-free number for questions or other requests. Calls will continue until contact (including voice mail) is made, a survey is returned, or about six weeks have elapsed. Upon request, we will mail a new questionnaire to the physician, fax the physician's office the questionnaire or send it by email.

Third Mailing

About 6-8 weeks after the end of follow-up calls, a third survey package will be sent to physicians who have not yet responded. Another round of follow-up calls will take place following the third mailing. Westat's experience has been that waiting 2 months to give the sample time to "rest" before re-approaching physicians will result in an increase in the response rate by at least a few percentage points.

Estimation Procedures

Sample Weighting

The sample weights will be formed in two stages reflecting both the probability of selection of a physician as well as an adjustment for nonresponse. These will be described in the following two sections on the development of the base weight (the reciprocal of the probability of selection) and the adjustments for nonresponse.

Base Weight Development

The probability of selection of the physician will be known at the time of sample selection. A sampled physician's base weight is simply the reciprocal of the chance of selection of the physician for the study.

Adjusting for Nonresponse

Nonresponse will be incurred both at the screener level and at the survey level. Adjustments will be undertaken to account for both.

Adjusting the Base Weight for Screener Nonresponse

Screener nonresponse will be incurred if a sampled physician cannot be located or if the physician's practice declines to participate in the study, in essence on behalf of the physician. There are a limited number of variables available for adjusting the weights for screener nonresponse, mainly those used for the sorting of the sample frame: specialty of sampled physician, census region, state, MSA status, sole proprietorship status, and, perhaps, gender are the main candidates. An evaluation of these variables will be undertaken to identify those that appear most effective in distinguishing between subgroups with different propensities to respond. We plan to employ CHAID for this purpose. CHAID is a commonly used tree-based algorithm for studying the relationship between a dependent variable and a series of predictor variables.

Once the CHAID analysis is completed, cells will be formed from the variables identified as effective in distinguishing between response propensities, and the weights of participating physicians associated with the cell will be adjusted to compensate for those in the same cell who do not participate.

Adjusting the Weights of Screener Respondents to account Survey Nonresponse

Information collected as part of the screener may serve to distinguish sampled physicians in terms of response propensity. Variables representing this information can be incorporated into a CHAID analysis to form cells for the purpose of adjusting the weights of screener participants for nonresponse to the survey instrument. "Size of practice" and "type of practice" (HMO, single office, etc.) are two possible variables potentially useful for nonresponse adjustment purposes.

Variance Estimation

For purposes of variance estimation and analyses, replicate weights will be created using a jackknife replication methodology.

Expected Levels of Precision

The targeted number of responding physicians for the study is 600. The expected margin of error for an estimated percentage of 50 percent at the 95 percent confidence level with an effective sample size of 600 (since no design effects will be incurred due to clustering or differential sampling rates) is roughly plus or minus 4 percent. The expected margin of error for 300 respondents (approximately the number expected for each specialty separately) is about plus or minus 5.7 percent.

3. Methods to Maximize Response Rates and Deal with Nonresponse

The data collection procedures described above were developed to maximize response rates. Screener nonresponse will be incurred if a sampled physician cannot be located or if the physician's practice declines to participate in the study, in essence on behalf of the physician. Survey nonresponse will arise if a screener has been completed but the survey is not completed by the sampled physician.

Some tracing will take place prior to the start of the field period and is intended to provide the most accurate starting sample possible. Tracing as needed throughout the field period is designed to ensure that all sampled respondents are contacted and offered an opportunity to participate. Once a physician is located, determined through screener data, repeated mailings are made, as needed, with the intent to motivate respondents to participate in the survey. A reminder in a different mode (the reminder phone call) is designed to capture the attention of respondents who have not responded to mailed invitations. Additionally, respondents will be mailed a \$50 incentive with the invitation to complete the survey to try to maximize response rates.

As described above, adjustments will be made to the sample weights to account for nonresponse to the screener and survey.

Nonresponse Bias Analyses

Because the response rate for this study is expected to be lower than 80 percent, we plan to undertake nonresponse bias analyses to help assess the extent to which there is potential for nonresponse bias to arise in the survey estimates. The process of adjusting the weights for screener nonresponse serves as one vehicle for these analyses. Subgroups or adjustment cells where screener response rates are particularly low can be readily identified through CHAID analyses, indicating potential sources of bias. The corresponding CHAID analysis to establish

cells for nonresponse adjustment among completed screeners provides another such vehicle for these analyses.

4. Tests of Procedures or Methods to be Undertaken

The draft survey has been reviewed by a 13 person Technical Advisory Group that included experts from HHS (CDC, AHRQ and HRSA), academia, the medical, and the nonprofit sector. In addition, nine cognitive interviews by telephone were conducted with primary care physicians who treat adult patients to pretest the revised draft. The cognitive interview participants were purposively recruited from primary care medical practices with differing numbers of physicians and different types of practice ownership (e.g., physician/ physician group, HMOs, hospital-owned practices, government-owned clinics, and clinics affiliated with medical schools). Recruiting techniques included requests for assistance from members of practice-based research groups and large health care systems, the distribution of recruiting flyers, and referrals from early participants to other eligible participants. The survey included in **Attachment 8** reflects changes resulting from the analysis of the cognitive interview findings. We do not anticipate any additional survey changes other than possible changes resulting from OMB review.

5. Individuals Consulted on Statistical Aspects and the Individuals

Mr. Ralph DiGaetano is the statistical consultant for the survey. Westat, Inc. will be overseeing the data collection and topline analysis of the Oral Health Survey. These efforts will be supervised by Mr. DiGaetano:

Ralph DiGaetano, MA
Senior Statistician, Westat, Inc.
301-294-2062
RalphDigaetano@westat.com

Ms. Martha Franklin is the survey design consultant for the survey:

Martha Franklin, MA
Instrument Design Specialist, Westat, Inc.
301-610-4826
MarthaFranklin@westat.com