NONSAMPLING ERROR IN A SURVEY OF DEPARTMENT OF DEFENSE HEALTH CARE BENEFICIARIES

Nancy A. Clusen, Eric Schone, Daniel Kasprzyk, Mathematica Policy Research
Thomas V. Williams, Department of Defense
Nancy A. Clusen, 600 Maryland Ave., SW, Suite 550, Washington, DC 20024, nclusen@mathematica-mpr.com

Key Words: Noncontact, Nonresponse, Mode effect, Error of measurement, Error analysis (Mathematics)

1. Introduction

This article reviews three sources of nonsampling error in the Health Care Survey of Department of Defense Beneficiaries (HCSDB): errors associated with the sampling frame, error due to unit nonresponse, and error due to measurement issues associated with the mode of data collection. The HCSDB, conducted in 2000, included an experiment in which nonrespondents were contacted by telephone in an effort to understand, identify, and quantify the bias due to nonresponse. The results of this experiment provide some insight into the extent of nonsampling error in the survey. This paper describes the errors, their cause, and their impact on survey estimates. We also offer some suggestion for reducing nonsampling error in the HCSDB.

2. Survey Design

2.1 Survey Objectives

The Adult HCSDB is the primary tool with which the TRICARE Management Activity (TMA) of the Assistant Secretary of Defense (Health Affairs) monitors the opinions and experiences of military health system (MHS) beneficiaries. The HCSBD was conducted annually between 1995 and 2000. Since 2001, the HCSDB has been conducted quarterly. The HCSDB is designed to provide information on the following questions:

- How do MHS beneficiaries rate their health care, their health plan, and primary care manager (PCM)?
- Do beneficiaries experience problems accessing care for themselves or their children?
- Do MHS beneficiaries experience difficulties in dealing with their health plan for claims processing of customer service issues?
- Do beneficiaries' ratings of their health plan, health care and primary care manager change over time?
- Is health care at military and civilian treatment facilities (MTFs and CTFs) meeting TRICARE standards?
- Is the level of use of preventive health care services consistent with national goals, such as those outlined in Healthy People 2010?

2.2 Sampling Frame

The target population for the HCSDB is all adults eligible to receive military health care benefits. The sampling frame includes all beneficiaries eligible for the survey as of a given reference date for the quarter. To be eligible for the HCSDB, individuals must be:

- Eighteen years of age or older on the reference date
- Eligible for military health care benefits as of the reference date
- Not incapacitated, incarcerated, or deceased
- The beneficiary or sponsor of the beneficiary must have been a member of one of the following: Army, Navy, Air Force, Marine Corps, Coast Guard, Public Health Service (PHS), or National Oceanic and Atmospheric Administration (NOAA)
- The beneficiary or sponsor of the beneficiary must have been one of the following: active duty, recalled to active duty, academy student/Navy OCS, National Guard, Reserve, transitional loss (RIF), or retired

The DoD Defense Manpower Data Center (DMDC) prepares the sampling frame, which consists of selected variables for each MHS beneficiary in the Defense Enrollment Eligibility Reporting System (DEERS) as of a specified reference date. DEERS includes everyone who is eligible for MHS benefits. DEERS includes those on active duty, those retired from military careers, immediate family members and surviving family members of those in the previous two categories. Using this frame, the sample was drawn independently within strata using a permanent random number (PRN) technique (Ohlsson 1995; Creel et al.

2002). In the 2001 HCSDB 45,000 adult beneficiaries were selected for each quarterly survey.

2.3 Mode of Data Collection

The HCSDB is currently a quarterly mail survey of a sample of MHS beneficiaries. The HCSDB mailing process is designed so each beneficiary with a useable address receives up to four mailings: a notification letter, a questionnaire, a reminder or thank you postcard, and a second questionnaire. The first questionnaire mailing is referred to as the first wave mailing and the second questionnaire mailing is the second wave mailing. There is no follow-up of nonrespondents after the second wave mailing.

3. An Experiment to Measure Nonsampling Error

In 2000, TMA contracted with MPR to conduct a study to evaluate the effectiveness of four follow-up approaches for contacting nonrespondents from the 2000 HCSDB mail survey. The experiment had one objective of interest here: to determine whether nonresponse bias affected survey results on access to and satisfaction with health care. Nonresponding beneficiaries after the first wave were randomly assigned to a treatment, which varied the time of contact, the type of request, and the length of the contact. In each of the four treatment groups, nonrespondents were contacted by telephone and asked to complete the questionnaire or an abbreviated questionnaire (Clusen and Schone 2000). However, like many studies of this type, we were not able to obtain 100 percent response from the sample of nonrespondents. Therefore, the measurement of nonresponse bias is also subject to nonresponse bias. Although the focus of the experiment concerned measuring nonresponse bias, the study provided significant insight into issues associated with the sampling frame and the mode of data collection. The discussion below is based on the results of the study.

3.1 Measuring Nonsampling Error: Errors Associated with the Frame

Noncontact in a mail survey can occur several ways, and the researcher can control but a few of them. The questionnaire must arrive at the correct address, the household "gate keeper" must forward the questionnaire to the sampled beneficiary, and the beneficiary must open the envelope. The events that occur within the household are beyond a researcher's direct control, although an enticing package for the questionnaire may help (Dillman 1999). However, mailing the questionnaire to the correct address is clearly the responsibility of the researcher and the failure to do so contributes to nonresponse (Lessler and Kalsbeek 1992).

3.1.1 Errors Associated with the Frame: Source of Address

The sampling frame for the HCSDB is based on the population data file constructed from DEERS, a data system that includes all MHS beneficiaries as of the survey reference date. The file contains variables required for sampling and data collection, such as mailing information and other locating information. Before data collection, this file is compared to the National Change of Address (NCOA) database to obtain updated address information. For the first mailing, we use the address provided by NCOA. Records not updated by the NCOA vendor are mailed to one of the three DEERS addresses: residential address, beneficiary's sponsor address, and military unit address. For mailing purposes, the residential address was given preference over the beneficiary's sponsor address. Likewise, the beneficiary's sponsor address is given preference over the military unit address.

Updating addresses is a continuous process throughout the data collection period. During data collection, address updates are obtained from three sources: (1) self-report by beneficiaries; (2) address correction information from the United States Postal Service (USPS); and (3) materials returned by the USPS as non-deliverable.

Based on the four quarters of the 2001 HCSDB, 6,121 of 180,000 sampled beneficiaries did not have sufficient address information and were not included in any of the mailings. Nearly 81 percent of the returned surveys were mailed to the beneficiary address supplied by DEERS. However, 67 percent of questionnaires were either not returned or were returned as non-deliverable. Most of these, 90 percent, were not returned (Mathematica Policy Research 2002). Previous research demonstrated that many of these questionnaires did not in fact reach the intended sampled beneficiary.

3.1.2 Errors Associated with the Frame: Incorrect Addresses

Evidence suggests that a significant portion of nonresponse in the HCSDB stems from the failure to make contact with the sampled beneficiary. In a 1999 study of nonresponse to the HCSDB, more than one-quarter of those interviewed by telephone reported not receiving the questionnaire (Bajaj et al. 1999). Furthermore, the fact that two-thirds of the beneficiaries selected for the 1999 study could not be matched to a working telephone number is further suggestive that incorrect addresses are responsible for much of the nonresponse.

The 2000 nonsampling error study provided us the means to estimate the proportion of the sample with incorrect addresses. To investigate nonresponse resulting from noncontact, we compared the response rates of beneficiaries whose address could and could

not be linked to a telephone number. We assumed if a telephone number was linked to a sampled beneficiary, then it was more likely the address was correct. Because the telephone nonresponse study was limited to beneficiaries in the continental United States, we exclude overseas beneficiaries from these analyses.

To measure the association of telephone numbers with correct addresses, we calculated three rates. First, from the sample of first wave nonrespondents with telephone numbers, we calculated the percent that eventually responded. This rate is 23 percent. Second, from the sample of first wave nonrespondents without telephone numbers, we calculated the percent that eventually responded. Only 15 percent responded, substantially lower than the proportion with telephone numbers. Third, based on information from the nonresponse telephone survey, 82 percent of the located telephone numbers were correct and working.

To estimate the percent of questionnaires mailed to incorrect addresses, we made two assumptions. First we assumed a correct telephone number meant a correct address. Consequently, we assumed that 82 percent of beneficiaries with a telephone number have a correct address. Second, we assumed all beneficiaries with valid addresses had the same second wave response rate.

Table 1: Calculation of Valid Address Rate

First wave respondents (1)	32%
First wave nonrespondents (2)	68
Of which: Nonrespondents with no telephone number (3)	65
Of which: Response rate (4)	15
Of (2): Nonrespondents with telephone number (5)	35
Of which: Response rate (6)	23
Of (5): Nonresponse with valid telephone number (7)	82
Valid address second wave response rate (8) = (6) / (7)	29
Estimated valid addresses with no telephone numbers (9) = (4) / (8)	53
Estimated nonrespondents with valid address (10) = (3)(9) + (7)(5)	63
Full sample estimated valid addresses (11) = (10)(2) + (1)	75

As Shown in Table 1, based on the first assumption, the estimated second wave response rate for valid addresses is the ratio of the percent of nonrespondents with telephone numbers that eventually respond (23 percent) to the percent of nonrespondents with a valid telephone number and therefore a correct address (82 percent). As a result, an estimated 29 percent of beneficiaries with a valid address will respond to the second wave of data collection. In later calculations we use this second wave response rate for all beneficiaries with a valid address, regardless of whether or not we have a telephone number.

Assuming that all beneficiaries with valid addresses had a second wave response rate of 29 percent, the percent of beneficiaries with no telephone number whose address was correct is the ratio of the percent of nonrespondents without telephone numbers that eventually respond (15 percent) to the percent of beneficiaries with a valid address that respond in the second wave. Therefore, 53 percent of beneficiaries with no telephone number have a correct address.

We then calculated among first wave nonrespondents the proportion of beneficiaries with a correct address. In the nonresponse telephone survey, about 65 percent of the sampled nonrespondents could not be matched to a telephone number. Therefore, the proportion of first wave nonrespondents with a correct address and no telephone number is 34 percent (53 percent of 65 percent). Of the 35 percent that could be matched to a telephone number, only 82 percent were valid numbers. Therefore, the proportion of nonrespondents with a correct address and a telephone number is 29 percent (82 percent of 35 percent). As a result we estimated 63 percent of first wave nonrespondents.

Using this measure, 37 percent of unreturned first wave questionnaires, or more than one-third of the nonrespondents, resulted from an incorrect address. For the full sample, the incorrect address rate is estimated to be somewhat less, 25 percent. However, it appears that a significant source of nonresponse results from the inability to contact the sampled beneficiary.

3.2 Errors Associated with Nonresponse

Survey estimates based on respondent data only may be biased with respect to describing characteristics of the population. This error is composed of two elements, the nonresponse rate and the difference between respondents and nonrespondents:

$$y_r = y_n + \left(\frac{nr}{n}\right)\left(y_r - y_{nr}\right)$$

where y_r is the statistic estimated from the r respondent cases; y_n is the statistic estimated for all n sample cases; and y_{nr} is the statistic estimated from the nr nonrespondent cases.

Therefore, nonresponse error is due to the proportion not responding to the survey and the difference of the statistic calculated between respondents and nonrespondents. This expression can be applied to both unit nonresponse—a sampled beneficiary does not respond to the survey—or item nonresponse—a particular questionnaire item is not

answered (Groves 1989). In this paper we discuss only unit nonresponse in the HCSDB.

3.2.1 Errors Associated with Nonresponse: Unit Nonresponse

Historically, response rates to the HCSDB have been below 50 percent (see Figure 1). In 1999, the HCSDB response rate was 35 percent, as compared to 50 percent the year previous. The decline prompted an investigation of the characteristics of 1999 nonrespondents (Bajaj et al. 1999). In 2000, we continued our study of nonrespondents with an experiment to evaluate the effectiveness of four follow-up approaches for contacting nonrespondents (Clusen and Schone 2000). The results of this study allow us to describe the characteristics of nonrespondents, particularly those with incorrect addresses, and to estimate the nonresponse bias in selected survey estimates.

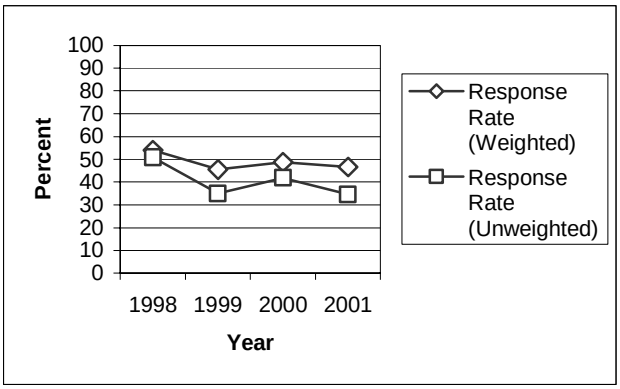

Figure 1: HCSDB Response Rates Over Time

3.2.2 Errors Associated with Nonresponse: Unit Nonrespondents

This study found that respondents and nonrespondents differ on several characteristics. Using information from the sample frame, we found:

- Nonrespondents are more likely to be young (less than 39 years old)
- Nonrespondents are more likely to serve in active duty or be a family member of someone serving in active duty
- Nonrespondents are more likely to be enlisted personnel

These differences are true for early nonrespondents (nonrespondents after the first wave) or later nonrespondents (nonrespondents after the second wave).

3.2.3 Errors Associated with Nonresponse: Nonlocated Cases

Because noncontact has been shown to contribute significantly to nonresponse, we divided nonrespondents into two group, located and nonlocated. Recall that beneficiaries for whom we were able to identify a telephone number were more likely to be located than beneficiaries who did not have a telephone number. Therefore, finding a beneficiary telephone number was used as a proxy for locating the correct address. Again, using information from the sample frame, as compared to located nonresponding beneficiaries, beneficiaries who were not located have the following characteristics:

- Nonlocated beneficiaries are more likely to be young (less than 29 years old)
- Nonlocated beneficiaries are more likely to serve in active duty or be a family member of someone serving in active duty
- Nonlocated beneficiaries are more likely to be enlisted personnel
- Nonlocated beneficiaries are more likely to be male

The differences in the demographic variables between nonrespondents with correct addresses and nonrespondents without correct addresses and the less-than-full response to the nonresponse follow-up suggests that the estimates of nonresponse bias are also subject to nonresponse error. However, this study allows us some limited ability to measure survey estimates of nonrespondents.

3.2.4 Errors Associated with Nonresponse: Comparing Respondents and Nonrespondents

Recall that in each of the four treatment groups, nonrespondents were contacted by telephone and asked to complete the questionnaire or an abbreviated questionnaire. Given the significant differences in demographic characteristics between respondents and nonrespondents in the treatment groups, we suspected that survey estimates are biased. Table 2 shows significant differences between the respondents and nonrespondents. Some differences are observed in "factual" rather than "opinion" type questions. Because the nonrespondent group is disproportionately enrolled in TRICARE Prime, estimates based on the respondent group only underestimate the percent of beneficiaries enrolled in TRICARE Prime and the percent of beneficiaries using Prime most often. Furthermore, it appears that estimates based on respondents only overestimate utilization (percent of beneficiaries visiting a specialist, percent with a personal physician) and underestimate health plan ratings and health status.

Table 2: Comparison of Estimates, Respondents and Nonrespondents

	Respondents	*Nonrespondents*	*Difference*
% Enrolled in Prime	31.8	52.5	-20.7*
% Using Prime Most	30.0	43.7	-13.7*
% with Personal Doctor	71.3	63.5	7.8*
% Rating Personal Doctor 8 or Higher	76.3	82.2	-5.9*
% Visited a Specialist	25.5	13.5	12.0*
% Rating All Health Care 8 or Higher	70.3	77.9	-7.6*
% Rating Health Plan 8 or Higher	54.9	62.6	-7.7*
% with No Doctor's Visits to Military Facility	60.7	53.1	7.6*
% Never Waiting for Well-Patient Visits	71.1	71.0	0.1
% Never Waiting More than 30 Minutes	37.8	48.1	-10.3*
% Rating General Health Excellent	14.1	22.7	-8.6*

*p < 0.05

3.3 Error Associated with the Mode of Data Collection

Other research has found that health care plan beneficiaries offer more positive ratings over the telephone as compared to self-administered questionnaires (Fowler, Gallagher, and Medereend, 1999). Moreover, previous research has shown that respondents offer more positive health assessments to telephone interviewers than to self-administered questionnaires (Hochstim 1967). Because we designed the follow-up to include telephone data collection, we can evaluate how the mode of data collection affects survey estimates. We investigated whether the change in the mode of data collection from mail to telephone could explain the apparent differences in estimates especially the differences for health plan rating and health assessment.

Our ability to test mode effects is limited by the study design. We did not randomly assign beneficiaries to receive a mailed questionnaire or a telephone interview. However, we were able to use cross tabulations and models to evaluate the impact of the mode. We expected that some of the differences in the estimates between mail respondents and telephone respondents were due to differences in demographic composition of the two groups. However, when we crossed treatment groups by demographic variables, we still saw differences in health plan and health care ratings. After controlling for beneficiary demographic characteristics, beneficiaries interviewed by telephone rated their health plan and health care significantly higher than did mail respondents. The percent of beneficiaries who rate both health care and health plan an 8 or higher (on a scale of 1 to 10) were 5 to 10 percentage points higher for each group of telephone respondents as compared to mail respondents.

To test further whether or not the data collection mode affects health care and health plan ratings, we constructed models predicting ratings by beneficiary age, sex, active duty status, TRICARE enrollment status, and treatment by mode. As can be seen in Table 3, beneficiaries interviewed by telephone have actual ratings significantly higher than the predicted ratings. Moreover, because beneficiaries interviewed by telephone are younger and are more often active duty, characteristics associated with lower ratings of health plans and health care, differences adjusted for beneficiary characteristics are larger than unadjusted differences.

Table 3: Results of Predictive Model

	Health Care Rating		*Health Plan Rating*		*Percent in Excellent Health*	
	Predicted	Actual	Predicted	Actual	Predicted	Actual
Treatment 1: Telephone	7.73	8.33*	6.91	7.34*	17	23*
Treatment 1: Mail	7.89	7.74	7.04	7.12	15	17
Treatment 2: Mail	7.81	7.98	6.94	6.92	16	18
Treatment 3: Telephone	7.67	8.23*	6.82	7.67*	18	23
Treatment 4: Telephone	NA	NA	6.94	7.38*	17	23
All Telephone	7.70	8.23*	6.90	7.47*	17	23*
All Mail	7.84	7.93	7.02	7.13	16	17

*$p < 0.05$

We also investigated whether self-reported health status was higher for telephone respondents as compared to mail respondents. We found some evidence the mode of data collection does affect the distribution of responses. As is the case for plan ratings, adjusted health status ratings are also higher for telephone respondents. In this case the evidence for a mode effect is weaker. We see that the percent reporting excellent health is significantly higher than predicted for only one of three telephone treatment groups and for telephone respondents overall.

4. DISCUSSION

The HCSDB is an important tool because it offers TMA timely data concerning access to, use of, and satisfaction with the military health system. However, survey estimates appear to be subject to non-ignorable nonsampling errors. Problems with the sampling frame contribute to nonresponse, and survey nonresponse results in nonresponse error.

To address the frame error problems, we might consider experimenting with the use of the military unit address for active duty beneficiaries or try alternative algorithms for rotating between residential, beneficiary's sponsor, and military unit address. To address the problem of nonresponse error, we might consider an experiment with express mailings or telephone follow-up. Fowler et al. found that a telephone follow-up of mail nonrespondents to the Consumer Assessment of Health Plans (CAHPS) survey could produce less biased samples than a mail-only methodology (2002). However, as we have seen, telephone follow-up may have unintended consequences on the measurement of health assessment and health plan and health care ratings. Nonsampling error occurs in all sample surveys. Having taken steps to identify and quantify selected sources of nonsampling error, the challenge for the HCSBD is to make incremental changes to reduce nonsampling error in the face of limited resources.

5. REFERENCES

Bajaj, R., Maxfield, M., Sargent, K.D., Schone, E., Werner, P., and White, K. (1999). *Investigation of Non-response in the 1998 Health Care Survey of DoD Beneficiaries*. Washington, DC: Mathematica Policy Research.

Clusen, N. and Schone, E. (2000). *Investigation of Nonresponse in the 1999 Health Care Survey of DoD Beneficiaries.* Washington, DC: Mathematica Policy Research.

Creel, D., Jang, D., and Williams, T.V. (2002) "Permanent Random Number Technique to Minimize Response Burden in Repeated Surveys." *2002 Proceedings of the American Statistical Association, Survey Research Methods Section* [CD-ROM]. Alexandria, VA: American Statistical Association.

Dillman, D. (1999). *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.

Fowler, F.J, Jr., Gallagher, P.M., and Mederend, S. (1999). "Comparing Telephone and Mail Response to the CAHPS™ Survey Instrument." *Medical Care*, 37, MS41-49.

Fowler, F.J, Jr., Gallagher, P.M., Stringfellow, V.L., Zaslavsky, A.M., Thompson, J.W., and Cleary, P.D. (2002). "Using Telephone Interviews to Reduce Nonresponse Bias to Mail Surveys of Health Plan Members." *Medical Care*, 40, 3, 190-200.

Groves, R.M. (1989). *Survey Errors and Survey Costs.* New York: Wiley.

Hochstim, J.R. (1967). "A Critical Comparison of Three Strategies of Collecting Data from Households." *Journal of the American Statistical Association*, 62, 976-989.

Lessler, J.T. and Kalsbeek, W.D. (1992). *Nonsampling Error in Surveys*. New York: Wiley.

Mathematica Policy Research. (2002) *2000 Health Care Survey of DoD Beneficiaries Adult Technical Manual*. Washington, DC: Mathematica Policy Research.

Ohlsson, E. (1995). "Coordination of Samples Using Permanent Random Numbers." In *Business Survey Methods* edited by B.G. Cox et al. New York: Wiley.