

## **THE EFFECT OF SURVEY FOLLOW-UP ON NONRESPONSE BIAS: 2002-2003 JOINT CANADA / UNITED STATES SURVEY OF HEALTH**

Stephen J. Blumberg<sup>1</sup>, Karen E. Davis, Meena Khare, Michael E. Martinez

### **ABSTRACT**

To achieve greater response rates for the U.S. sample, the 2002-2003 Joint Canada/United States Survey of Health fielded additional telephone calls and refusal conversion attempts for three months following the planned close of data collection. Final weighted survey estimates were compared with estimates from the U.S. National Health Interview Survey and with weighted estimates based on interviews completed prior to the planned close of data collection, interviews completed with no more than 10 dials, and interviews completed without refusal conversion efforts. Nonresponse bias remained constant despite efforts to increase response rates.

KEY WORDS: Nonresponse bias; Refusal conversion efforts, Response rates, Telephone survey.

### **RÉSUMÉ**

Pour atteindre des taux de réponse plus élevés pour l'échantillon américain dans le sondage conjoint Canada/États-Unis sur la santé, on a effectué des appels téléphoniques additionnels et d'autres essais pour convertir des refus pour les trois mois suivant la date prévue d'arrêt de collecte de données. Les estimations pondérées finales de sondage ont été comparées avec les estimations du U.S. National Health Interview Survey et avec les estimations pondérées basées sur les interviews complétés avant la date prévue d'arrêt de collecte, les interviews complétées avec au plus 10 appels et les interviews complétées sans efforts pour convertir les refus de réponse. Le biais de non réponse est resté constant malgré les efforts pour augmenter le taux de réponse.

MOTS CLÉS : Biais de non réponse; efforts pour convertir les refus de réponse; sondage téléphoniques; taux de réponse.

## **1. INTRODUCTION**

### **1.1 Background**

In the United States, response rates for random-digit-dial telephone surveys have either declined or have been sustained only through increasing levels of effort and resources (Curtin, Presser, and Singer 2005). The downward pressure on response rates is the result of many influences, including an increasing reluctance to participate in surveys and an increasing use of technologies to avoid unwanted telephone calls. The general reluctance to participate in surveys arises in part from increasing concerns about privacy and the security of personal information. New telephone screening technologies, including answering machines, caller-ID, and privacy managers, have become increasingly popular due to the increase in unsolicited calls from telemarketers and charitable organizations.

The telephone-based 2002-2003 Joint Canada / United States Survey of Health (JCUSH) was not immune to these pressures, and the response rate from the survey suffered as a result. On March 31, 2003, when the data collection effort was scheduled to end, the overall response rate for the United States sample was 45.3%. Because the response rate is often used as one of the key measures of survey quality, Statistics Canada and the U.S. National Center for Health Statistics, who jointly sponsored the survey, decided to increase their efforts to complete additional interviews and raise the response rate following the scheduled close of data collection.

---

<sup>1</sup> Affiliation for all authors: National Center for Health Statistics, Centers for Disease Control and Prevention, 3311 Toledo Road, Hyattsville, Maryland, USA, 20782. Address e-mail to Stephen J. Blumberg, Ph.D., [sblumberg@cdc.gov](mailto:sblumberg@cdc.gov).

The response rate, however, is only a measure of the *potential* for nonresponse bias from a survey. Nonresponse bias is a function of both the response rate and the characteristics of the nonresponders (Groves and Couper 1998). That is, the degree of bias resulting from the failure to contact or interview some eligible persons will be high if the number of nonrespondents is large *and* if the nonrespondents differ in a meaningful way from the respondents, but it will be low if either the number of nonrespondents is small or if the nonrespondents do not differ in a meaningful way from the respondents. Several recent studies have found examples where the difference between respondents and nonrespondents was sufficiently small that nonresponse bias was not affected when the response rates were in the range of 40% to 70% (Curtin, Presser, and Singer 2005; Keeter et al. 2000).

The purpose of this paper is to examine whether the increased efforts to complete interviews in the United States reduced bias and therefore increased the quality of estimates from the JCUSH.

## **1.2 The Joint Canada / United States Survey of Health**

The JCUSH was a collaborative effort between the U.S. National Center for Health Statistics (of the U.S. Centers for Disease Control and Prevention) and Statistics Canada to collect health information from both Canadian and U.S. residents using the same questionnaire and survey procedures (Sanmartin et al. 2004). Questionnaire items were drawn from ongoing health surveys in both countries, and comparable sample design, administration, and weighting procedures were employed. All interviews were conducted using random-digit-dialling, with telephone calls to both countries originating from call centers at Statistics Canada. The goal of this comparability was to improve the accuracy of cross-national comparisons on measures of health status and access to health care services.

The target population for the JCUSH was the civilian noninstitutionalized resident population of adults aged 18 years or older in Canada and the United States. Random-digit-dialling was used to identify a representative sample of households with eligible adults, and one adult was randomly selected from each household. Because one additional goal of the JCUSH was to produce reliable estimates for adults aged 65 years and older, the probability of selection of adults in this age range was increased. Interviews were conducted in a choice of English or French for Canadian residents and English or Spanish for U.S. residents.

The data collection for the JCUSH began on November 4, 2002, and concluded on March 31, 2003 for Canadian residents and on July 14, 2003 for U.S. residents. The overall response rate for Canadian residents was 65.5%, whereas the overall response rate for U.S. residents was 50.2%. The lower rate in the U.S. was primarily a function of a lower interview completion rate; the proportion of known households within which an interview was completed was 62.4% in the U.S., compared with 90.9% in Canada.

Additional details about the survey design are available elsewhere (Statistics Canada and U.S. National Center for Health Statistics 2004). This paper is limited to data from JCUSH/US, the U.S. resident portion of the sample.

## **1.3 Survey Follow-Up Efforts**

To minimize nonresponse, interviewers were instructed to make all reasonable effort to complete interviews, including the scheduling of appointments to complete interviews at times that were convenient for the respondents, numerous callbacks at various days and times to reach households, and scripted responses to avert possible refusals by providing additional information about the survey. In addition, prior to the first contact by the interviewer, advance letters explaining the purpose of the survey and the confidentiality of responses were mailed to households for which a mailing address was available.

Nevertheless, response rates for the JCUSH/US remained low. In March 2003, follow-up letters were mailed to households that had refused when first invited to participate in the survey. These letters stressed the importance of the survey and the household's cooperation. This was followed by another call from a senior interviewer, a project supervisor, or another interviewer trained to convey the importance of the survey to potential respondents. In April 2003, follow-up letters were also mailed to households with telephone numbers that continued to ring with no answer or that always were answered by answering machines. These letters were followed by additional calls to these telephone numbers. In June 2003, additional calls were made to households that had twice previously refused to participate.

If calling had stopped on March 31, 2003 (as originally planned), only 4,672 interviews would have been completed with U.S. residents. Instead, by July 14, 2003, a total of 5,183 interviews had been completed (an addition of 511 completed interviews). The median number of calls necessary for a completed interview was 6, with a range of 1 to 49. Some telephone numbers received 50 or more calls without achieving a completed interview.

If efforts to complete an interview had been stopped after 10 calls to each telephone number, only 3,818 interviews would have been completed. This loss of 1,365 interviews, or 26.3% of all completed interviews, would have resulted in a response rate of 37.0%.

If efforts to complete an interview had been stopped after 15 calls to each telephone number, 4,470 interviews would have been completed. This loss of 713 interviews, or 13.8% of all completed interviews, would have resulted in a response rate of 43.3%.

Finally, if all further calls had been ceased following the first time a household respondent had refused to participate, 4,314 interviews would have been completed. This loss of 869 interviews, or 16.8% of all completed interviews, would have resulted in a response rate of 41.8%.

## 2. METHOD

### 2.1 Overview

As noted earlier, the purpose of this paper is to examine whether the increased efforts to complete additional interviews would have reduced bias and therefore increased the quality of estimates for the U.S. population. To answer this question, four nonindependent subsets of the data were identified: a) Interviews completed with 10 or fewer calls; b) interviews completed with 15 or fewer calls; c) interviews completed on or before March 31, 2003; and d) interviews completed without the household initially refusing to participate. Each of these subsets represents the sample that would have been available had various rules been in place to determine when to stop calling telephone numbers.

Each subset was reweighted to permit estimates that generalize to the U.S. adult population aged 18 years or older. These estimates were then compared with final estimates from the JCUSH/US and with 2002 estimates from the U.S. National Health Interview Survey (NHIS; Lethbridge-Cejku, Schiller, and Bernadel 2004). The NHIS is an annual multistage probability-based household survey that collects comprehensive health-related information from a large sample of households representing the civilian noninstitutionalized household population of the U.S. This face-to-face survey interview is conducted by the U.S. National Center for Health Statistics, is administered by trained field representatives from the U.S. Census Bureau, achieves high response rates (88.1% in 2002), and provides high quality data that are considered one of the best sources for estimates of health status and health care access in the U.S.

### 2.2 Weighting Procedure

In order for estimates produced from survey data to be representative of the target population, and not just the sample, a survey weight is assigned to each person for whom an interview is completed. This weight corresponds to the number of persons represented by the respondent in the target population. Sampling weights for U.S. residents who participated in the JCUSH began as the inverse probability of the telephone number being selected from the universe of all telephone numbers in the U.S. This initial weight was adjusted for household nonresponse, multiple telephone lines per household, the probability of selection of an adult from all adults in the household, and nonresponse among selected adults. Finally, post-stratification adjustments were used to ensure that population totals from the JCUSH/US match population control totals from an independent, more reliable source. For the JCUSH/US, the independent source was the October 2002 Current Population Survey (CPS), conducted by the U.S. Census Bureau. Poststratification adjusted the JCUSH/US weights so that demographic distribution from the JCUSH/US would match the distributions of the population by sex, race/ethnicity (Hispanic, non-Hispanic black, and other non-Hispanic adults), and age (18-34, 35-44, 45-54, 55-64, and 65+ for non-Hispanic non-black adults, and 18-44, 45-64, and 65+ for others) from the CPS.

For each subset of the data, new weights were computed for this analysis by recalculating the final post-stratification adjustment factors. A flag was used to identify whether the respondent was excluded from the subset ( $flag_i = 1$ ) or not ( $flag_i = 0$ ). The new weight was then formulated as:

$$(1) \quad Wt\_flag_i = \begin{cases} 0 & \text{if } flag_i = 1, \\ \left[ \frac{\sum_{j=pstrat\_cell} fwgt_j}{\sum_{j=pstrat\_cell} fwgt_j I(flag_i)} \right] \cdot fwgt & \text{if } flag_i = 0. \end{cases}$$

where  $fwgt_j$  is the final JCUSH/US weight based on all 5,183 completed interviews,  $pstrat\_cell$  is each of the 17 poststratification cells, and  $I(\ )$  denotes the indicator function. Note that  $I(flag_i = 0) = 1$  and  $I(flag_i = 1) = 0$ . Table 1 displays several basic statistical measures for the weights of each subset. The last column provides the final JCUSH/US weight based on all 5,183 completed interviews for comparison. Note that the medians are higher for the subset weights due to the smaller sample sizes. Table 1 also displays several basic statistical measures for the post-stratification adjustment factors for each subset.

### 3. RESULTS AND CONCLUSION

#### 3.1 Overall Bias

The final two columns of Tables 2 and 3 permit an examination of the bias that exists between the JCUSH/US and the NHIS. Relative to the NHIS, JCUSH/US estimates reveal fewer adults without a high school diploma, fewer adults with at least some college education, and more adults with only a high school diploma. The JCUSH/US data also reveal fewer low-income adults and more insured adults less than 65 years of age, but more adults who could not afford needed prescription medications within the past year. Moreover, the JCUSH/US data reveal fewer adults with excellent or very good health status, more adults with fair or poor health status, and more adults who are unable to work due to their health. In addition, the JCUSH/US data reveal fewer adults ever diagnosed with hypertension and fewer adults aged 65 years or greater with limitations of activities. The magnitude of the bias was generally less than five percentage points.

#### 3.2 Impact of Survey Follow-up on Bias

A look at the sociodemographic distribution that would have resulted from the JCUSH/US if certain stopping rules had been in place generally reveals little impact on the magnitude of the bias. Estimates of the population distribution by sex,

**Table 1 – Statistical Summary Measures for Sample Weights and Adjustment Factors for Post-Stratification Cells**

Summary Measure	Based on	Based on	Based on	Based on	Based on All Completed Interviews (n = 5,183)
	Interviews Completed with 10 or Fewer Calls (n = 3,818)	Interviews Completed with 15 or Fewer Calls (n = 4,470)	Interviews Completed On or Before March 31, 2003 (n = 4,672)	Interviews Completed Without a Prior Refusal (n = 4,314)	
<b>Sample weights</b>					
Minimum	8,338.6	6,987.4	6,876.7	7,326.0	6,342.3
Median	48,505.5	43,774.3	44,305.3	48,182.4	38,735.8
Maximum	364,719.9	324,811.3	260,210.2	259,690.8	235,011.5
Coefficient of variation	59.34	58.21	55.22	53.04	55.16
Standard deviation	32,081	26,882	24,398	25,380	21,967
<b>Adjustment factors for post-stratification cells</b>					
Minimum	1.149	1.035	1.000	1.035	--
Median	1.379	1.169	1.114	1.202	--
Maximum	1.809	1.572	1.323	1.349	--
Coefficient of variation	9.55	6.83	3.21	6.85	--
Standard deviation	0.130	0.079	0.036	0.082	--

age, and race/ethnicity would remain generally unchanged because these distributions were imposed during the post-stratification of the weights. (Minor differences may exist as a result of rounding.) Table 2 reveals that estimates for the proportion of U.S. adults with incomes below USD\$20,000 would have been less biased if less effort had been used to complete interviews (i.e., if calls had ended after 10 attempts), but that this reduction would have been small (i.e., bias of 4.0 percentage points instead of 4.6). Similar results were revealed for educational attainment: The estimate of the proportion of U.S. adults with either a college degree or some college education would have been less biased if less effort had been used to complete interviews (in this case, if refusal conversion had not been used), but this reduction would have been small (i.e., bias of 1.9 percentage points instead of 2.6).

Table 3 presents similar data for measures of health status and health care access. For most of these measures, however, estimates would have become slightly more biased had less effort been used to complete interviews. Stopping calls after only 10 attempts would have resulted in the greatest increase in bias for these measures, although relative to the bias resulting from the other stopping rules, the increase would have been quite small (i.e., generally less than one percentage point). It is worth noting, however, that the other three stopping rules examined would have increased bias by no more than 0.5 percentage points for any measure, with one exception (limitation of activity).

### 3.3 A Look at the Characteristics of Hard-to-Reach Respondents

As noted previously, these stopping rules would have had no impact on the weighted estimates of the population distribution by sex, age, and race/ethnicity because these distributions were imposed during the post-stratification of the weights. However, unweighted statistics from each subset reveal the sample distribution by sex, age, and race/ethnicity

**Table 2 – Percent Distributions of the U.S. Population by Selected Sociodemographic Characteristics**

Characteristic	2002-2003 Joint Canada / United States Survey of Health					2002 U. S. National Health Interview Survey
	Based on Interviews Completed with 10 or Fewer Calls	Based on Interviews Completed with 15 or Fewer Calls	Based on Interviews Completed On or Before March 31, 2003	Based on Interviews Completed Without a Prior Refusal	Based on All Completed Interviews	
Sex						
Male	48.0	48.0	48.0	48.0	48.0	48.0
Female	52.0	52.0	52.0	52.0	52.0	52.0
Age						
18-44 years	52.3	52.3	52.3	52.3	52.3	52.5
45-64 years	31.7	31.7	31.7	31.7	31.7	31.4
65+ years	16.0	16.0	16.0	16.0	16.0	16.1
Race / Ethnicity						
Hispanic, any race	11.6	11.6	11.6	11.5	11.6	11.0
Non-Hispanic, white only	69.8	69.9	69.9	70.0	70.0	73.2
Non-Hispanic, black only	11.6	11.7	11.7	11.7	11.8	11.4
Non-Hispanic, other	7.0	6.8	6.8	6.8	6.5	4.4
Income						
< USD\$20,000	15.4	15.1	15.0	14.9	14.8	19.4
>= USD\$20,000	84.6	84.9	85.0	85.1	85.2	80.6
Education						
Less than high school	11.5	11.4	11.7	11.7	11.8	16.6
High school graduate	37.7	37.7	37.2	36.7	37.4	29.9
Some college	14.3	14.2	13.9	14.0	14.0	20.0
College graduate	36.5	36.6	37.1	37.6	36.9	33.5
Geographic region						
Northeast	19.9	20.4	19.6	18.6	20.1	19.3
Midwest	22.9	22.4	22.4	22.7	22.1	24.4
South	35.1	35.2	34.7	35.0	35.1	37.0
West	22.1	22.0	23.2	23.8	22.7	19.3

**Table 3 –Estimated Percent of U.S. Residents with Selected Health Characteristics**

Characteristic	2002-2003 Joint Canada / United States Survey of Health					2002 U. S. National Health Interview Survey
	Based on Interviews Completed with 10 or Fewer Calls	Based on Interviews Completed with 15 or Fewer Calls	Based on Interviews Completed On or Before March 31, 2003	Based on Interviews Completed Without a Prior Refusal	Based on All Completed Interviews	
Health status						
Excellent or very good	58.4	58.8	58.8	58.8	58.9	62.3
Good	26.6	26.8	26.8	26.9	26.6	25.5
Fair or poor	15.0	14.4	14.4	14.3	14.5	12.2
Ever diagnosed with hypertension	22.7	22.9	22.6	22.7	22.7	24.3
Ever smoked 100 or more cigarettes	47.0	46.2	46.0	45.9	45.7	45.1
Has health insurance <sup>1</sup>	89.1	88.9	88.6	88.6	88.6	84.5
Had unmet need for prescription medication, past 12 months	10.6	10.3	10.0	9.8	9.9	7.2
Has limitation of activity <sup>2</sup>	47.5	47.6	51.4	49.5	48.5	61.5
Unable to work due to health	9.3	8.8	8.5	8.4	8.5	7.2

<sup>1</sup>Limited to adults aged 18-64 years

<sup>2</sup>Limited to adults aged 65 years or older

that would have resulted from these stopping rules. These sample distributions shed some light on the characteristics of adults living in households with telephones who were difficult to reach but eventually agreed to participate in the survey.

As revealed in Table 4, women were overrepresented in the sample based on all completed interviews, suggesting that women were more likely than men to participate in the JCUSH/US. The overrepresentation of women would have been even greater if less effort had been used to complete interviews; the increased effort resulted in more participation among men.

Because of the oversample of adults aged 65 years or older, these adults were overrepresented in the final sample and in all subsets resulting from the imposition of the stopping rules. The overrepresentation of these older adults would have been even greater if calls had stopped after 10 or 15 attempts, but would have been less pronounced if calls had stopped after a single refusal. This result suggests that older U.S. residents were more likely than adults from other age groups to initially refuse participation, but also more likely to be reached and to participate when contacted again. For younger adults (18-44 years), the data reveal that the converse is true: More calls were required to reach these younger adults, but once contacted, they were less likely to refuse to participate.

Table 4 also reveals that non-Hispanic black adults and (to a lesser extent) Hispanic adults were underrepresented in the sample based on all completed interviews, suggesting that adults in these minority groups were generally less likely than non-Hispanic white adults to participate in the JCUSH/US. Moreover, non-Hispanic black adults and Hispanic adults were harder to reach, and therefore their underrepresentation would have been even greater if less effort had been used to complete interviews.

### 3.4 Summary and Conclusion

The response rate for the U.S. resident sample in JCUSH was low, in part because of Americans' general reluctance to participate in surveys and because of their increasing use of new technologies to avoid unwanted telephone calls (Curtin, Presser, and Singer 2005). This low response rate raised concerns about the potential for nonresponse bias, and indeed some bias was revealed when estimates from JCUSH/US were compared with estimates from the NHIS. An examination of the estimates that would have been obtained if less effort had been used to complete interviews demonstrated that, for most measures, the magnitude of the bias would have been either unchanged or only slightly greater.

**Table 4 – Percent Distributions of the JCUSH/US Sample by Selected Sociodemographic Characteristics**

Characteristic	2002-2003 Joint Canada / United States Survey of Health				
	Based on Interviews Completed with 10 or Fewer Calls	Based on Interviews Completed with 15 or Fewer Calls	Based on Interviews Completed On or Before March 31, 2003	Based on Interviews Completed Without a Prior Refusal	Based on All Completed Interviews
Sex					
Male	42.3	42.3	42.5	42.7	42.9
Female	57.7	57.7	57.5	57.3	57.1
Age					
18-44 years	42.3	43.6	45.5	47.1	45.2
45-64 years	32.3	32.6	32.5	32.5	32.6
65+ years	25.4	23.8	22.0	20.4	22.2
Race / Ethnicity					
Hispanic, any race	9.5	9.7	10.8	11.5	10.6
Non-Hispanic, white only	76.4	76.2	75.0	73.7	75.1
Non-Hispanic, black only	7.4	7.5	7.7	8.1	7.9
Non-Hispanic, other	6.7	6.6	6.5	6.6	6.4

The increased effort to complete interviews also increased the costs of the survey. Was this effort worth the investment of time and resources, given that the final response rate was still only 50.2%? The added effort was perhaps not worth the price if the goal of such effort was to substantially reduce the bias that may result from nonresponse. On the other hand, the added effort may be considered cost-effective if the goal of such effort was to increase the representation of men, young adults, and minority groups in the survey sample.

#### REFERENCES

- Curtin, R., Presser, S., & Singer, E. (2005). "Changes in telephone survey nonresponse over the past quarter century". *Public Opinion Quarterly*, **69**, 87-98.
- Groves, R., & Couper, M. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley and Sons.
- Keeter, S., Miller, C., Kohut, A., Groves, R., & Presser, S. (2000). "Consequences of reducing nonresponse in a large national telephone surveys". *Public Opinion Quarterly*, **64**, 125-48.
- Lethbridge-Cejku, M., Schiller, J. S., & Bernadel, L. (2005). "Summary health statistics for U.S. adults: National Health Interview Survey, 2002". National Center for Health Statistics. *Vital and Health Statistics*, **10**(221).
- Sanmartin, C., et al. (2004). *Joint Canada / United States Survey of Health, 2002-03* (Catalogue 82M0022-XIE). Analytic report released jointly by Statistics Canada and the National Center for Health Statistics. Ottawa: Statistics Canada. Available at [http://www.cdc.gov/nchs/data/nhis/jcush\\_analyticalreport.pdf](http://www.cdc.gov/nchs/data/nhis/jcush_analyticalreport.pdf).
- Statistics Canada & U.S. National Center for Health Statistics (2004). *Joint Canada / United States Survey of Health Public Use Microdata File User Guide*. Ottawa: Statistics Canada. Available at [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Dataset\\_Documentation/JCUSH/questionnaires/English\\_Userguide.pdf](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/JCUSH/questionnaires/English_Userguide.pdf).