

**Validation of Socially Desirable Health Behaviors and
Implications for Case-Control and Other Health Surveys:
Does Salience Matter? ***

Diane O'Rourke¹, Timothy Johnson¹, Jane Burris¹, and Richard Warnecke²
University of Illinois at Chicago

¹ Survey Research Laboratory, ² Cancer Center and Health Policy and Research Centers

ABSTRACT

Many major health surveys ask respondents about their health behaviors, including the use of cancer prevention tests for women, such as mammograms and Pap smear tests. In addition, case-control studies related to cancer query women with the diseases as well as control cases to compare them to. An important question is: Does the salience of the topic, including past negative experiences, affect the quality of the recall? If it does, and if "cases" recall more accurately than "controls," the variability in the quality of the data may significantly affect the value of the conclusions.

We present findings from a factorial experiment designed to evaluate four approaches to collecting autobiographical health information that was then compared to medical records. Data were collected from a random sample of women aged 50 and older who were either interviewed by telephone or by using audio computer-assisted self-interviewing equipment in their homes. They were asked for the dates they received mammograms and Pap smear tests over the past 36 months. Upon obtaining written permission, medical records staff at their provider facilities recorded the actual dates of such tests and the responses were compared. Of the 1,005 women who participated in the study, medical record information was obtained for 588. The accuracy of the information was analyzed controlling for measures of salience, including having had a "problem" mammogram (requiring a followup mammogram or biopsy), having breast cancer, having a relative with cancer, particularly breast cancer, or having had a "problem" Pap smear test (requiring re-testing or followup). Results indicate that there was no significant difference in the accuracy of the recall by the salience measures.

Therefore, at least for common cancer prevention tests such as mammograms and Pap smears that are taken frequently and easily recalled, past experience and possible salience is not an issue.

INTRODUCTION

Health promotion and disease prevention are goals of most nations (1). Many major health surveys ask respondents about their past and current health behaviors. In addition to the response errors that may occur in all survey research (2-5), recall of autobiographical information is even more problematic. Over the past two decades, researchers studying the cognitive processes of survey response have gained a better understanding of the process and how to control error (6-13), but information about health behaviors, particularly socially desirable or undesirable ones, and particularly dating those behaviors are still very problematic.

For recall of autobiographical events, memory is basically a function of time (between the recollection and the event) and the salience of the event. In this paper, we focus on salience, which has been defined as being "a function of the unusualness of an event, its economic and social costs and benefits, and its continuing consequences" (14, p.88). Salient events may be more likely to be retained in long-term memory and may be more likely to be accessible to survey respondents (14, 15). Thus, in addition to memory problems that relate to overall autobiographical recall, another potential error factor may be the differential salience of the topic among respondents. This is particularly important for case-control studies that query women with a disease/condition as well as control cases (who do not have that disease/condition) to compare them to. An important question is: Does the salience of the topic, including past negative experiences, affect the quality of the recall? If it does, and if "cases" recall more accurately than "controls," the variability in the quality of the data may significantly affect the value of the conclusions in many epidemiological surveys

Paper presented at the 57th Annual Conference of the American Association for Public Opinion Research, May 13-16, 2004, Phoenix, Arizona.

*This study was supported by National Cancer Institute grant RO1-CA77315 awarded to our friend and colleague, the late Seymour Sudman, Principal Investigator, and authors Johnson, O'Rourke, and Warnecke, co-Investigators

HYPOTHESIS

The research hypothesis is that salience, as measured by multiple variables, will improve the accuracy of the reporting of health behaviors.

MATERIALS AND METHODS

The data for this presentation come from a study that was designed to see if changes in data collection and/or question order could improve the accuracy of recall about the mammograms and Pap smear tests women had received in the past three years. Responses were validated by checking medical records. (Results of that study, as well as greater detail on the methods, are described in Johnson et al. 2003 AAPOR presentation (16).)

Sample design

While most studies of health behaviors utilize samples from Health Maintenance Organizations (HMOs) to simplify the validation process, this study began with a city-wide random-digit-dialing screening for eligible respondents - women aged 50 and older. This approach was economically feasible because the community, Champaign-Urbana, Illinois, is served primarily by two large health care providers.

A random-digit-dial sample of 12,923 telephone numbers was used to screen households. All women 50 years and older, who had lived and received health care in the Champaign-Urbana area for at least three years prior to the date of their interview, were eligible to participate in the study. Households with African-American women were over-sampled. All households were screened by telephone and then either interviewed by telephone (CATI) or by audio computer-assisted self interview (ACASI). A total of 1,005 interviews were completed across both conditions: 790 via telephone and 215 were self-administered via computer. The response rate for the CATI interviews, estimated using American Association for Public Opinion Research (AAPOR)(57) standard formulas (RR3) was 56.1 percent. For the ACASI condition, two response rates were calculated, one for CATI screening and the other for the interview. The telephone screening response rate for the subsequent ACASI cases was 33.1 percent. The response rate for the screened ACASI households was 78.6 percent; the overall AAPOR response rate for this condition was 26.6 percent. (See Table 1 for a complete breakdown of rates.) Of the women completing interviews, 82.8 percent consented to having their medical records abstracted ($n = 832$). Seventy-five percent of those women ($n = 621$) granting consent actually returned signed consent forms to our offices. Of these, record abstractions

were collected for 588 cases (94.7%). While most physicians outside of the two large group practices filled out requests for record checks for a small number of patients, a few either refused or did not return information. All results presented below are based on 588 cases.

Study procedures

During the screening process, interviewers explained the details of the study, answered respondent questions, and gained cooperation. If an eligible respondent was identified and agreed to participate in the CATI condition, the interview occurred immediately following the screening. In the ACASI condition, screened households were re-contacted by trained face-to-face interviewers. The questionnaire was completed in the homes of the respondents and at their convenience. Face-to-face interviewers presented themselves equipped with a laptop computer, a set of headphones, and relevant project materials.

All interviews were conducted between October 2001 and April 2002 in English only by female interviewers. The average telephone interview was approximately 25 minutes in length while the average ACASI version was approximately 35 minutes. Medical record abstraction was conducted by records personnel from each of the respective medical facilities. The only data extracted from respondents' medical records were the dates of their physical or gynecological exams, Pap smear tests, and mammograms.

Measures

We focused on self-reports of physical exams, mammograms, and Pap smear tests, but for this presentation we will exclude information on physical exams since they are not particularly salient to the respondents. The self-report indicator was based on reports of these procedures during a three-year period prior to the date of interview. The self-reports were then compared to objective medical records. Using these records as the standard, we calculated a variety of agreement measures. For this analysis, we looked at concordance and accuracy. Concordance is the percentage of cases in which the experience or nonexperience of the Pap or mammogram was accurately reported, overall, within the three-year period. For accuracy, a scale was constructed to indicate the number of months, plus or minus, the self report deviated from the record, with 0 being totally accurate (i.e., both the self report and the record noted the same month and year). A positive number would indicate forward telescoping

(reporting the procedure more recently than indicated in the record) and a negative number would indicate backward telescoping.

The questionnaire contained several variables that could relate to the salience of the mammogram and Pap smear test. They include: the self-reported health status of the respondent, the number of health problems she had had in the past three years, whether she had had a breast-related problem (including cancer) in the past three years, whether “a problem” had ever been found after a mammogram, and whether a mammogram had been followed up by another mammogram or a biopsy. In addition, we asked whether any blood relatives had had cancer and, in particular, breast cancer. For Pap smear tests, we asked whether there had been a problem with any Pap tests and if any Paps had to be followed up with subsequent Paps or additional testing. Affirmative response to any of these queries could indicate an increase in the salience of these topics for the respondent that could affect the accuracy of her responses. Respondents were grouped as to whether any responses indicated a “mammogram problem” or not or indicated a “Pap problem” or not. A combined measure of having either a Pap or mammogram problem was also constructed. The appendix lists the questions and answers used to form those variables. The following analyses are based on these groupings.

Analysis

We compared the dates of mammograms and Pap tests as indicated by the respondents and as shown in their medical records. According to our research hypothesis, it is anticipated that higher salience of the topic would increase the concordance and accuracy of the self report, as verified by the medical record.

We initially conducted univariate and bivariate analyses of the concordance rates and accuracy for Pap smears and mammograms by the measures of salience we developed, as well as relevant sociodemographic measures, including age, race, and educational achievement.

Logistic and OLS regression were subsequently employed to explore the independent effects of these variables on report-record concordance. Odds ratios (OR) and 95 percent confidence intervals (CI) are reported for each logistic regression model.

RESULTS

The mean age of respondents was 63.4 years (standard deviation; SD = 10.2; range = 50-94). Six percent reported less than a high school education, 22.8 percent were high school graduates, 25.2 percent had completed some college work, 15.6

percent were college graduates, and 30.4 percent had completed graduate training. This latter finding is reflective of the fact that Urbana-Champaign is home to a large public university. The sample was primarily Caucasian (85.8%), with 10.9 percent African American and 3.2 percent representing other race/ethnic categories. Most reported having some form of health insurance (97.4%) and 58.4 percent reported having a family member who had had cancer.

The three-year self-reported prevalence of having received Pap smears and mammograms in this sample were 60.7 percent and 81.9 percent, respectively. The three-year prevalence for each procedure, based on information available in the medical records, was 51.0 percent and 75.0 percent, respectively.

Table 2 shows the concordance measures (percent matched) for women with and without Pap and mammogram “problems” and with and without family members with cancer and specifically breast cancer. In five of the seven comparisons shown, the difference was in the direction hypothesized; that is, a higher percentage of women in the salient groups correctly matched the procedures. However, none were statistically significant, although one (women with either a Pap problem or a mammogram problem) was significant at the .07 level.

Table 3 shows the results of a logistic regression of concordance on the salience variables and demographic variables. None were independently predictive of self-report/medical record concordance.

Analyses were also conducted using “accuracy” as the dependent variable (plus or minus months of self report from medical record date). Both Pap and mammography screening were reported fairly accurately by this population. The mean difference between the dates reported for Pap tests and the dates found in the medical records was 0.14 months (SD = 2.75), a value not significantly different from zero ($t = 0.81$, $df = 258$, ns). The white respondents’ mean accuracy for Pap tests was .0225 months (where 0 is a perfect month’s match) and African American respondents accuracy was 1.667 months (forward telescoping about a month and a half). Again, there were few significant differences when examining our salience measures. In an OLS regression model that examined predictors of Pap smear reporting accuracy, which included overall health status rating, total number of health problems, Pap problems, family members with cancer, education, age, and race, the only significant finding was race. For the three-category racial variable (white, African American, and other), African American women were significantly more

likely (at the .02 level) to forward telescope than the white women.

The mean difference between self-reported date of most recent mammography screening and the date found in the medical records was 0.26 months ($SD = 3.66$; $t = 1.43$, $df = 416$, ns). When examined in a logistic regression model, none of the salience or sociodemographic variables examined were significantly associated with mammography reporting accuracy. When accuracy was defined as a self-report that was within one month (plus or minus) of the medical record date, no significant differences were found for either Paps or mammograms on any of the salience or sociodemographic variables.

DISCUSSION

Salience, or the importance of a topic to a person, has long been shown to affect unit nonresponse as well as item nonresponse (15). The question studied here was whether it would significantly affect accurate recall of the dates of previous Pap smear tests and mammograms. In this instance the answer is no. Our analyses looked at three levels of precision: (a) whether a medical record and self-report both reported or did not report a procedure in the past three years, (b) whether the self-reported date was plus or minus one month from the recorded date, and (c) how many months from the reported date was the self-reported date. Salience was measured using questionnaire variables such as having had problems with mammograms and Paps, having family members with cancer, and indicating the number of current health problems and self-reported current health status. While some of the results indicated trends in the hypothesized direction, few were statistically significant. In addition, there were few significant differences in accuracy of recall by demographic differences such as age, education, and race. There could be several reasons for these findings.

For one, the overall level of accuracy of the self-reported dates was very high. Overall concordance rates were .79 for Pap smear tests and .88 for mammograms. That leaves very little room for improvement or for group differences on salience. Second, the two cancer prevention procedures are often done annually on a well-scheduled basis, so recalling a date that may occur every 12 months is not a great cognitive burden for most respondents. Previous research has demonstrated higher accuracy in the reporting of cancer screening tests among persons who receive them on a regular basis (17).

Third, because cancer-prevention testing is now so common, it may not differentiate well between

groups. It is possible that salience would be a more important variable for health behaviors that are not as frequent or periodic or less unique. On the other hand, in the example of a case-control study, a "case," who has more medical events to recall will have more chances for error than a "control" person who does not. "No event" is easier to recall than specific events.

Fourth, the quality of the salience measures employed in this study might also be questioned. We attempted to evaluate salience indirectly using three approaches: (1) current health status, (2) negative personal health experiences, and (3) family history of cancer. We reasoned that persons rating their health as poorer, those having personally experienced negative health events, and those with family members who had experienced cancer would all interpret personal cancer screening experiences as more salient and hence be more likely to be able to accurately recover information about them from long term memory. Using indirect measures of salience, this hypothesis was not confirmed. More direct approaches to measuring salience should be explored. For example, it might be useful to ask respondents how important their health and the health of relatives are to them.

Two limitations to this study could also have affected the results. Medical records were used as the "gold standard" for evaluating the quality of the self-reports. Some past studies have determined that these records can be incomplete and not totally reliable (18-22). This is particularly the case for procedures, such as mammograms and Pap tests, which are evaluated outside of the doctor's office and must be entered separately into the patient's record.

Another limitation is the low response rate of the ACASI sample. This is because after being screened by telephone, only half (46%) of the respondents agreed to a face-to-face interview. This was not totally surprising with potential respondents who were all women aged 50 and over, many living alone, even for a survey that was sponsored by the local university. Nonetheless, this study's use of a random community sample should be considered an advantage.

While nonsignificant findings are typically not very helpful for adding to a body of literature, in this case they are at least reassuring. That is, in the case of health behaviors that are common for both "cases" and "controls," there does not seem to be differential error in the reporting of the events, even if the outcomes of those events were more negative for some respondents or were related to negative familial experiences, such as cancer. Further research should investigate the effects of salience on

health behaviors and events that are less common and/or frequently experienced.

REFERENCES

1. McGinnis J. Health objectives for the nation. *American Psychologist* 1991; 46:520-4.
2. Biemer PP, Groves RM, Lyberg LE, Mathiowetz NA, Sudman S. *Measurement Errors in Surveys*. New York: John Wiley & Sons, 1991.
3. Groves RM. *Survey Errors and Survey Costs*. New York: John Wiley & Sons, 1989.
4. Lyberg L, Biemer P, Collins M, de Leeuw E, Dippo C, Schwarz N, Trewin D. *Survey Measurement and Process Quality*. New York: John Wiley & Sons, 1997.
5. Sudman S, Bradburn NM. *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine, 1974.
6. Blair E, Burton S. Cognitive processes used by survey respondents in answering behavioral frequency questions. *Journal of Consumer Research* 1987; 14:280-288.
7. Jobe JB, Mingay DJ. Cognition and survey measurement: History and overview. *Applied Cognitive Psychology* 1991; 5:175-92.
8. Lessler JT, Tourangeau R, Sarter W. Questionnaire design in the cognitive research laboratory. *National Center for Health Statistics, 1989. Vital Health Statistics* 6(1).
9. Schwarz N, Sudman S. *Autobiographical Memory and the Validity of Retrospective Reports*. New York: Springer-Verlag, 1994.
10. Sirken M, Herrmann D, Schechter S, Schwarz N, Tanur J, Tourangeau R. *Cognition and Survey Research*. New York: Wiley, 1999.
11. Sudman S, Bradburn NM, Schwarz N. *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey Bass, 1996.
12. Tanur JM. *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation, 1992.
13. Tourangeau R, Rips LJ, Rasinski, K. *The Psychology of Survey Response*. New York: Cambridge University Press, 2000.
14. Pearson RW, Ross M, Dawes RM. Personal recall and the limits of retrospective questions in surveys. Pp. 65-94 in JM Tanur (Ed.) *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation, 1992.
15. Sudman S, Bradburn N. *Asking Questions*. San Francisco: Jossey-Bass, 1982.
16. Johnson T, O'Rourke D, Burriss J, Warnecke R. Improving the validity of self-reports of cancer screening behaviors. Paper presented at the 2003 AAPOR conference.
17. O'Rourke D, Johnson T, Sudman S, Warnecke R, Jobe J. The effect of regularity on the accuracy of reporting medical tests. *American Statistical Association 1993 Proceedings of the Section on Survey Research Methods: Vol. I [AAPOR section]*, 324-329.
18. Demlo LK, Campbell PM, Brown SS. Reliability of information abstracted from patients' medical records. *Medical Care* 1978; 16: 995-1005.
19. Feigl P, Glaefke G, Ford L, et al. Studying patterns of cancer care: How useful is the medical record? *American Journal of Public Health* 1988; 78:526-533.
20. Horwitz RI, Yu EC. Assessing the reliability of epidemiologic data obtained from medical records. *Journal of Chronic Diseases* 1984; 37: 825-831.
21. Kosecoff J, Fink A, Brook RH, Chassin MR. The appropriateness of using a medical procedure: Is information in the medical record valid? *Medical Care* 1987; 25: 196-201.
22. Fomm, F., & Putnam, S. The validity of the medical record. *Medical Care* 1981; 19:310-315.

Table 1. Sample Rates: CATI and ACASI Screening and Interviewing

	CATI		ACASI				Combined	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate
Total Sample	8,699		4,224		268 ^a		4,224	
Non duplicates	8,698	100.0	4,202	99.5	264	98.5	4,202	99.5
Working numbers	7,348	84.5	3,511	83.6	263	99.6	3,510	83.5
Residential numbers	6,370	86.7	3,067	87.4	263	100.0	3,066	87.4
Contact to screener	5,309	83.3	2,519	82.1	258	98.1	2,513	82.0
Cooperation to screener	4,624	87.1	2,187	86.8	239	92.6	2,162	86.0
Eligible	1,057	22.9	587	26.8	236	98.7	559	25.9
Contact to final	1,005	95.1	571	97.3	232	98.3	539	96.4
Cooperation	790	78.6	264	46.2	204	87.9	204	37.8
Response rate ^b		56.1		33.1		78.6		26.6
Refusal rate ^c		23.9		47.1		10.8		52.2
Cooperation rate ^d		70.1		41.3		87.9		33.7

^a 4 duplicate cases screened

^b AAPOR response rate RR3

^c AAPOR refusal rate REF2

^d AAPOR cooperation rate COOP4

Table 2. Concordance of self report and medical record data by measures of salience

	Pap smear		Mammogram	
	%	(n)	%	(n)
Pap problem – yes	81.3	91		
Pap problem – no	78.7	497		
Mammogram problem – yes			90.6	127
Mammogram problem – no			87.2	460
Pap or mamm. problem – yes	77.0	196	91.3 ^a	196
Pap or mamm. problem – no	80.1	392	86.2	391
Family member with cancer				
- yes	80.9	340	86.7	339
- no	76.4	242	89.7	242
Family member with breast cancer				
- yes			90.6	85
- no			87.5	506

^a Pearson chi-square = 3.24, df = 1, asymp. significance = .072

Table 3. Logistic regressions of Pap smears and mammograms on salience and demographic variables

	Pap smear		Mammogram	
	<u>OR</u>	<u>95% CI</u> ^a	<u>OR</u>	<u>95% CI</u> ^a
Total health	0.827	0.648, 1.056	1.074	0.769, 1.500
Family	1.219	0.782, 1.902	0.747	0.425, 1.313
Education	1.010	0.875, 1.166	1.039	0.867, 1.246
African American	1.414	0.644, 3.105	0.594	0.268, 1.316
Other nonwhite race	0.727	0.189, 2.801	1.294	0.161, 10.419
Age	0.989	0.966, 1.012	0.992	0.965, 1.020
Health status rating	1.040	0.922, 1.173	1.087	0.937, 1.261
Pap problem	1.015	0.532, 1.933	-	-
Mammogram problem	-	-	1.344	0.609, 2.965

^a OR, odds ratio; CI, confidence interval

APPENDIX—Survey questions included as measures of salience

Total health: Total number of health problems in past three years

- In the past three years, have you had a problem with:
 - your heart, cholesterol, or hypertension?
 - diabetes or blood sugar?
 - a female or urinary problem?
 - a breast-related problem?
 - a cancer-related problem?

Health status rating: Sum of responses to two health status questions:

- Would you say that your health, in general, is excellent, very good, good, fair, or poor?
- Compared to other people your age, would you say that your health is (repeat)?

Family: Has any parent, sister, brother, or child of yours, related by blood, had cancer?

Pap problem: One or more Pap smear problems:

- For each Pap smear test mentioned, a “problem” response to the question:
 - Was everything OK or did they find a problem?
- For additional Pap smear tests, all answers except “part of a routine checkup”:
- Why did you get the additional Pap smear test? (Part of a routine checkup, because of symptoms, as a follow-up to previous problem, other)
- “Yes” answer to: Have you ever had additional tests because a Pap test came back positive, suggesting that there might be a problem?

Mammogram Problem: One or more mammogram problems:

- For each mammogram mentioned, a “problem” response to the question:
 - Was everything OK or did they find a problem?
- For additional mammograms, all answers except “part of a routine checkup”:
- Why did you get the additional mammogram? (Part of a routine checkup, because of symptoms, as a follow-up to previous problem, other)
- “Yes” answer to: Have you ever had an ultrasound or breast biopsy because something was found on a mammogram?