Published in final edited form as:

Med Care. 2013 October ; 51(10): 945–948. doi:10.1097/MLR.0b013e3182a5023d.

Do Additional Recontacts to Increase Response Rate Improve Physician Survey Data Quality?

Gordon B. Willis, PhD¹, Tenbroeck Smith, PhD², and Hyunshik J. Lee, PhD³

¹National Cancer Institute, 9609 Medical Center Drive, Rm 3E358, MSC 9762, Bethesda, MD, 20892-9762

²American Cancer Society, Inc, 250 Williams Street NW, Atlanta, GA 30303 Phone: (404) 327-6442; Fax: (404) 929-6832; Tenbroeck.Smith@cancer.org

³Westat, 1600 Research Blvd, Rockville, MD 20850 Phone: 301-610-5112; Fax: 301-294-2034; HyunshikLee@westat.com

Abstract

Background—Although response rates for physician surveys have been decreasing, it is not clear whether this trend is associated with an increase in survey non-response bias. One means for assessing potential bias is to conduct a level-of-effort analysis that compares data estimates for respondents interviewed during the first recruitment contact to respondents interviewed at later recontact cycles.

Methods—We compared early and later responders within the Survey of Physician Attitudes Regarding the Care of Cancer Survivors (SPARCCS) with respect to both demographic characteristics and aggregate survey responses to items on survivor care knowledge, attitudes and practices.

Results—Accumulating additional completions across each of four respondent contact attempts improved weighted response rates (35.0%, 46.9%, 52.3%, and 57.6%, respectively). However, the majority of estimates for analyzed variables remained relatively unchanged over additional cycles of recontact.

Conclusions—We conclude that additional respondent recontact attempts, especially beyond a single recontact, had little influence on key data distributions, suggesting that these were ineffective in reducing nonresponse bias. Further, the conduct of additional recruitment recontacts was an inefficient means for increasing statistical power. For the conduct of physician surveys, a practice that may in some cases be cost-effective, while also controlling total survey error, is to

Gordon Willis, PhD (corresponding author) National Cancer Institute Room 3E358 9609 Medical Center Drive, Rm 3E358, MSC 9762 Bethesda, MD 20892-9762 Phone: 240-276-6788 Fax: 240-276-7906 willisg@mail.nih.gov.

My co-authors and I are pleased to submit the revised manuscript: “Do Additional Recontacts to Increase Response Rate Improve Physician Survey Data Quality?” by Gordon Willis, Tenbroeck Smith, and Hyunshik Lee., for consideration as a Brief Report in Medical Care.

Each of the three authors has taken responsibility for (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article; and (3) providing approval of the version submitted for publication. None of the authors has any conflict of interest with respect to the research described; all authors have engaged in this research as part of their employment, respectively, at the National Cancer Institute, NIH; the American Cancer Society; and Westat.

Thank you, once again, for considering our manuscript.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

establish a larger initial sample; to either eliminate nonresponse follow-up or to limit this to one recontact; and to accept a somewhat lower final overall survey response rate.

INTRODUCTION

Physician surveys provide important data on patient care, health systems, and clinical practices, making them an essential element of health research and policy. However, surveying physicians is challenging, and physician surveys may suffer from lower response rates than surveys of the general population.^{1,2} Three systematic reviews of physician survey methods suggest ways to increase response rates and reduce survey error in physician surveys, based on the published literature.¹⁻³ These authors conclude that whereas several approaches for increasing response rate that are effective in the general populace also apply to physicians (e.g., financial incentives, special contacts, and personalization), there is insufficient evidence regarding the effectiveness of others – in particular, the optimum number of contact attempts within physician surveys.¹

These reviews have noted that even if additional recontacts of physicians result in increased response rates, this outcome does not necessarily represent an improvement in survey data quality, reiterating Groves'⁴ conclusion that response rate is not in itself a good measure of nonresponse bias. In practice, researchers often search for sources of potential bias in several ways.⁵ They may assess differences in demographic characteristics or in key data estimates between respondents and nonrespondents.³ Or, they can select a *level-of-effort* approach, which compares early and late responders to assess the effects of additional contact attempts on key measures. If additional contacts result in non-trivial changes in data distributions, then the steps necessary to collect those data are deemed worthwhile. In contrast, if the additional cases have little effect on substantive results, it can be argued that data collection could have been stopped earlier to reduce cost or shift resources elsewhere -- to increased sample size, pretesting of materials, or increasing timeliness and extent of the dissemination of results, etc., thereby maximizing *total survey quality*.⁶

Overall, past level-of-effort studies have found that additional recruitment contacts lead to few meaningful differences in key survey measures within physician surveys. Berk⁷ contrasted the data provided from groups of early, middle, and late physician responders, finding that only 3 of 14 statistical estimates studied changed by more than 5 percent, and only one by more than 10 percent. Berk did, however, note that accepting lower response rates reduces absolute sample size, thereby lowering statistical power, especially for subgroup analyses. Similarly, Schoenman, Berk, Feldman, and Singer⁸ assessed cumulative effects of increasing response rate from 50 to 65% through additional contact attempts and extension of the field period, for a telephone-administered physician survey. Again, differences in data estimates were mostly unaffected as later respondents were added. Guadagnoli and Cunningham⁹ and Sobal and Ferentz¹⁰ obtained similar findings, indicating that additional effort for purposes of maximizing response rate was generally not worthwhile. A review by Kellerman and Herold³ concluded that neither demographic nor practice characteristics of physicians differed between early and late respondents.

In the years since these studies were conducted, response rates have continued to decline for physician studies; McCleod, Klabunde, Willis, & Stark¹¹ report that for the period 2005-2009, only about a third of health care provider surveys reported response rates of at least 60%. Short of a methodological breakthrough, this trend may continue, potentially leading to a continuing decline in response rates. Given this potential threat to survey validity, for the current investigation we conducted a level-of-effort analysis to determine the effects of increasing sample size over several recontact cycles, for a more recent population-based survey of U.S. physicians. The fundamental research question is identical

to that posed by earlier studies cited above: Does additional level of effort result in meaningful changes in the data gathered? We also addressed a secondary research question: Whether or not estimates change significantly, do multiple recontacts contribute to sample size in such a way that statistical power is efficiently enhanced?

METHODS

We reanalyzed data from the Survey of Physician Attitudes Regarding the Care of Cancer Survivors (SPARCCS); see Potosky et al.¹² for details concerning sample, administration, and analysis. The American Medical Association (AMA) Physician Master File was used to obtain a nationally representative sample of oncologists and primary care physicians (PCPs) practicing family medicine, general internal medicine, or obstetrics/gynecology. Two separate but parallel versions of the questionnaire were created to enable comparisons between PCP and Oncologists. Screener telephone calls were first placed to the offices of sampled physicians to verify eligibility for participation and to obtain contact information. A total of 3,596 physicians received the questionnaires by mail in 2009, with a single \$50 incentive check. Screened physicians received up to four contact mailings, including three to the provider's office and one to a home address where possible, with reminder telephone calls after the 3rd and 4th mailings. For current purposes, we will consider each of the four successive mailing as a separate *contact cycle*, and refer to additional attempts to induce survey completion, beyond the initial contact attempt, as *recontacts*. Combined screener and survey response rates were calculated using the American Association of Public Opinion Research's (AAPOR) standard methods.¹³ The final weighted response rate using the AAPOR RR3 formula, which incorporates unscreened physicians with unknown eligibility, was 57.6% (unweighted 58.1%).

For the current level-of-effort analyses, a panel of cancer survivorship experts selected 30 critical demographic, knowledge, attitude, and practice-related variables from those analyzed by Potosky et al.¹² The critical research question was not whether the estimates at each contact cycle differ, but whether cumulative estimates obtained over subsequent cycles change as more cases are added.⁷ Therefore, frequency distributions were produced for each categorical variable, and mean values for each continuous variable, for each of the four contact cycles, in cumulative fashion. Hence, Cycle 4 is equivalent to the full sample, after four total contact attempts were made. Applying base weighting for probability sampling and nonresponse adjustment to reduce potential nonresponse bias, separate weights were computed for each of the four cycles, and all estimates and analyses applied these weights. The cumulative nature of the analysis led to large overlap in the four samples and therefore high inter-correlations, so analyses were conducted using a program (WesVar) that computes standard error estimation in a manner that accounts for this overlap using a replicate Jackknife method.¹⁴

We conducted a second level-of-effort analysis to investigate the extent to which additional recruitment over subsequent contact attempts enhanced statistical precision with respect to key sub-group comparison. Given that the comparison of PCPs and Oncologists was the primary aim of the SPARCCS survey, we compared these subgroups on all 30 targeted variables, and determined the number of PCP versus Oncologist differences that were significant ($p < .05$) for the full sample (Cycle 4), yet not at each earlier cycle.

Finally, in order to estimate potential cost savings for the SPARCCS survey associated with alternative recontact strategies, we applied mathematical cost modeling comparing the four-contact strategy to both one-contact and two-contact strategies (details are available in the Cost Analysis included as online Supplemental Digital Content).

RESULTS

Total sample sizes for the variables selected were: Full sample (at Cycle 4) = 2202; PCP = 1072; and Oncologist = 1130. Table 1 illustrates the size of the sample at each of Cycles 1-4. Overall, and also for each of the two physician subgroups, the four Cycles contained approximately 57%, 80%, 90%, and 100% of the final sample, respectively. Weighted cumulative response rates for each of the Cycles were 35.0%, 46.9%, 52.3%, and 57.6%. The first level-of-effort analysis (Table 2) illustrates the number of the 30 evaluated demographic and substantive survey variables that were statistically different between the full sample (Cycle 4) and each of the other Cycles. Even at Cycle 1, where one might expect to see considerable effects because that Cycle had accumulated only 57% of the eventual sample, only 3 of 30 estimates showed significant differences between Cycles 1 and 4; the largest of these differences was 3.4% in magnitude. Differences were fewer at Cycles 2 and 3, and similarly small in magnitude. Overall, additional respondent contact attempts led to minimal change in statistical estimates.

Even if successful follow-up contact attempts had little influence on estimates of key variables, we note that they did provide the benefit of increased sample size and hence power to obtain statistically significant effects. Overall, 20 of our 30 analyzed variables exhibited significant ($p < .05$) differences between PCPs and Oncologists at Cycle 4. Of these 20 items, 12 revealed significant differences at Cycle 1, 16 at Cycle 2, and 17 at Cycle 3. However, our mathematical cost model revealed that the final sample size obtained could have been achieved at lower cost through an alternative strategy that increased initial sample size and reduced number of recontacts. Comparing the four-contact strategy used in SPARCCS to a one-contact strategy (i.e., no follow-up) incorporating a larger initial sample, we estimate that the one-contact strategy would have reduced total cost by approximately 25% while achieving the same final sample size. Subsequent models suggested that a two-contact strategy (initial contact plus one recontact cycle), although more expensive than a single contact, would still have provided substantial savings, relative to four total contacts (approximately 15%).

DISCUSSION

Our analysis of SPARCCS data found few differences in estimates of physician demographic, knowledge, attitude, or practice variables between early and late responders, suggesting that resources expended to conduct additional contact attempts did little to reduce nonresponse bias. Additional recontact cycles were somewhat successful in increasing the survey sample size and enhancing our capacity for identifying significant PCP-ONC differences. However, we suggest that conduct of multiple cycles of recontact simply to increase power is likely to be inefficient. Instead, given the finding that multiple recontacts appear to reduce nonresponse bias minimally, **and that they inflate costs substantially**, a more efficient approach to achieving adequate power would be to incorporate a larger initial sample, and to either forego recontact entirely—or, as a compromise—to limit data collection to two total contacts. In particular, the selection of a larger initial sample, with fewer additional contacts, can be shown to reduce total survey costs whenever recontact cycles exhibit (a) equal or greater unit costs, and (b) lower cycle-specific response rates, relative to the initial contact (See Cost Analysis, Supplemental Digital Content). We note that studies incorporating expensive follow-up methods to convert hard-to-reach or recalcitrant respondents tend to drive up unit cost for later contact attempts significantly. Multiple contact attempts also require a longer survey field period, which typically increases time-dependent costs related to factors such as management personnel, survey processing systems, and maintenance of toll free numbers.

Overall, our findings echo Berk's⁷ observation that physician survey data quality is not clearly improved through intensive attempts to increase response rate to some critical level. In fact, results such as those reported here have been interpreted as suggesting that nonresponse bias is less of a concern in surveys of physicians than in general population surveys,^{1,3} perhaps because physicians as a group are more homogeneous than the general public with regard to demographic, knowledge and attitudinal characteristics. Another factor contributing to the stability of estimates we observed may be the effectiveness of nonresponse adjustment weighting for samples of physician respondents.

Limitations

As an important caveat, we stress that the results of our single case-study do not necessarily generalize to other physician surveys. Cumulative response rates in SPARCCS varied from 35.0% at Cycle 1 to 57.5% at Cycle 4, and it is unclear to what degree our results apply to surveys producing significantly different rates. Survey administration variables such as mode (e.g., mail versus Internet) may have major impact on selection of non-response follow-up strategy, as would analytic considerations such as the quality of information available for final sample weighting adjustment. Further, although issues of cost are vital to survey designers, we do not have access to data that would allow us to generally model cost savings associated with any particular approach.

Conclusions

Our findings suggest that attempts to maximize response rate through the devotion of significant resources to nonresponse follow-up may not be an optimal means for maximizing data quality for physician surveys having characteristics similar to the SPARCCS. Instead, practitioners should consider the tradeoffs associated with the alternative of starting with a relatively larger initial sample, and limiting resources devoted to recontact. Making such decisions depends on a consideration of relative cost, time, and effort associated with factors such as establishing the sampling frame, screening potential respondents for eligibility, balancing efforts to obtain completions across initial contact and recontact cycles, and obtaining information for use in re-weighting the final sample.

We note that our proposed approach runs contrary to a scientific culture which promotes response rate as the dominant measure of survey quality. That belief is reflected by scientific journals which favor manuscripts reporting response rates above a certain threshold (for example JAMA¹⁵ requests 60% or higher). Such rules, when coupled with inconsistent approaches to response rate calculation, may encourage authors to manipulate or at least to incorrectly compute response rates,¹¹ or to apply practices that increase response rate yet increase other types of error. A potential means for resolving this dilemma is to adopt the recommendations of Johnson and Wislar,⁵ who suggest that survey researchers conduct methodological studies – such as level-of-effort analysis – to assess nonresponse bias within sample surveys. Only by direct assessment of such sources of error can researchers determine the significance and implications of a particular response rate, and the overall level of quality of the design underlying the survey investigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

1. Field TS, Cadoret CA, Brown ML, et al. Surveying Physicians: Do Components of the “Total Design Approach” to Optimizing Survey Response Rates Apply to Physicians? *Med Care*. 2002; 40(7):596–605. [PubMed: 12142775]

2. VanGeest J, Johnson T, Welch V. Methodologies for Improving Response Rates in Surveys of Physicians: A Systematic Review. *Eval Health Prof.* 2007; 30:303–321. [PubMed: 17986667]
3. Kellerman S, Herold J. Physician response to surveys: A review of the literature. *Am J Prev Med.* 2001; 20(1):61–67. [PubMed: 11137777]
4. Groves R. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly.* 2006; 70(5):646–675.
5. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. *JAMA.* 2012; 307(17): 1805–6. [PubMed: 22550194]
6. Groves R, Lyberg L. Total survey error: Past, present, and future. *Public Opinion Quarterly.* 2010; 74(5):849–879.
7. Berk ML. Interviewing physicians: the effect of improved response rate. *Am J Public Health.* 1985; 75:1338–40. [PubMed: 4051077]
8. Schoenman JA, Berk ML, Feldman JJ, Singer A. Impact of differential response rates on the quality of data collected in the CTS physician survey. *Eval Health Professions.* 2003; 26(1):23–42.
9. Guadagnoli E, Cunningham S. The effects of nonresponse and late response on a survey of physician attitudes. *Eval Health Professions.* 1989; 12:318–328.
10. Sobal J, Ferentz K. Comparing physicians' response to the first and second mailings of a questionnaire. *Eval Health Professions.* 1989; 12:329–339.
11. McCleod CC, Klabunde CN, Willis GB, Stark D. Health care provider surveys in the U.S., 2000-2010: A review. *Eval Health Professions.* in press.
12. Potosky A, Han P, Rowland J, Klabunde C, Smith T, Aziz N, et al. Differences between primary care physicians' and oncologists' knowledge, attitudes and practices regarding the care of cancer survivors. *J Gen Intern Med.* 2011; 26(12):1403–10. [PubMed: 21785923]
13. American Association of Public Opinion Research. [August 26, 2012] Standard definitions: final dispositions of case codes and outcome rates for surveys. 7th ed.2011. Available at: http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions2&Template=/CM/ContentDisplay.cfm&ContentID=3156.
14. Brick JM, Morganstein D, Valliant R. Analysis of Complex Sample Data Using Replication. Available at: <http://www.westat.com/Westat/pdf/wesvar/ACS-Replication.pdf>.
15. JAMA. [August 26, 2012] Instructions for authors. Available at: <http://jama.jamanetwork.com/public/instructionsForAuthors.aspx>.

Table 1

Distribution of SPARCCS sample over Cycles 1-4, for All Physician, Primary Care Physicians (PCP), and Oncologists (total and percentage); and Weighted Response Rates.

Data collection Cycle (cumulative)				
	1	2	3	4
All physicians (n)	1251	1755	1976	2202
% of full sample	56.8%	79.7%	89.7%	100%
PCPs (n)	618	856	963	1072
% of PCPs	57.6%	79.8%	89.8%	100%
Oncologists (n)	633	899	1013	1130
% of Oncologists	56.0%	79.6%	89.6%	100%
Response rate ^I	35.0%	46.9%	52.3%	57.5%

^IWeighted response rates; unweighted were nearly identical.

Table 2

Number of statistically significant differences ($p < 0.05$) in estimates of 30 variables between Contact Cycle 4 (full sample) and each preceding Cycle for SPARCCS Survey: All physicians, Primary Care Physicians (PCP), and Oncologists.

Number (%) of significant differences between listed contact cycle, and Cycle 4 total sample			
Cycle	All Physicians	PCPs	Oncologists
1	3 (10.0%)	3 (10.0%)	3 (10.0%)
2	0 (0.0%)	0 (0.0%)	2 (6.7%)
3	2 (6.7%)	2 (6.7%)	0 (0.0%)