

© Health Research and Educational Trust
DOI: 10.1111/j.1475-6773.2010.01129.x
RESEARCH ARTICLE

A Randomized Trial Comparing Mail versus In-Office Distribution of the CAHPS Clinician and Group Survey

Michael P. Anastario, Hector P. Rodriguez, Patricia M. Gallagher, Paul D. Cleary, Dale Shaller, William H. Rogers, Karen Bogen, and Dana Gelb Safran

Objective. To assess the effect of survey distribution protocol (mail versus handout) on data quality and measurement of patient care experiences.

Data Sources/Study Setting. Multisite randomized trial of survey distribution protocols. Analytic sample included 2,477 patients of 15 clinicians at three practice sites in New York State.

Data Collection/Extraction Methods. Mail and handout distribution modes were alternated weekly at each site for 6 weeks.

Principal Findings. Handout protocols yielded an incomplete distribution rate (74 percent) and lower overall response rates (40 percent versus 58 percent) compared with mail. Handout distribution rates decreased over time and resulted in more favorable survey scores compared with mailed surveys. There were significant mode–physician interaction effects, indicating that data cannot simply be pooled and adjusted for mode.

Conclusions. In-office survey distribution has the potential to bias measurement and comparison of physicians and sites on patient care experiences. Incomplete distribution rates observed in-office, together with between-office differences in distribution rates and declining rates over time suggest staff may be burdened by the process and selective in their choice of patients. Further testing with a larger physician and site sample is important to definitively establish the potential role for in-office distribution in obtaining reliable, valid assessment of patient care experiences.

Key Words. Survey mode, survey distribution method, patient experiences, physician performance measurement

Patient experience measures are now recognized as central to a comprehensive assessment of health care quality. Large-scale initiatives such as pay-for-performance and public reporting programs require reliable and valid information about physician-level performance from probability samples of

adequate size (Safran et al. 2006; Rodriguez et al. 2009). The Agency for Healthcare Research and Quality's Consumer Assessment of Healthcare Providers and Systems (CAHPS) Clinician and Group survey (C/G CAHPS) has been approved by the National Quality Forum (NQF) as a measure of patient experiences with individual physicians and their practice sites (Agency for Healthcare Research and Quality, American Institutes for Research, Harvard Medical School, and RAND Corporation 2006).

A major issue when using surveys for large assessments is the cost of survey administration. When compared with mail or phone administration of surveys, in-office distribution of survey instruments ("handout") can reduce costs by relying on staff time and greatly limiting the costs associated with mail or telephone-based data collection (Schonlau, Fricker, and Elliott 2002; Gribble and Haupt 2005). Modes that require interaction with an interviewer, such as face-to-face or telephone administration, have been shown to elicit more positive ratings of care than mail surveys (Walker and Restuccia 1984; de Vries et al. 2005), and modes that reduce the elapsed time between the visit and survey completion have also been shown to produce more favorable assessments (Savage and Armstrong 1990; Kinnersley et al. 1996). There is limited evidence on differences in response rates and response patterns between surveys that are distributed in an office and mailed surveys. One study found that handing out surveys resulted in higher response rates, more favorable ratings, less overall variation in patient response, and higher item-level nonresponse compared with surveys administered by mail (Gribble and Haupt 2005). However, more information is needed about whether and how survey distribution method affects response rates and the results obtained.

We conducted a multisite randomized trial of handout and mail survey distribution to adult patients from the panels of 15 primary care physicians in a large multispecialty medical group in New York. We compared physician and

Address correspondence to Dana Gelb Safran, Sc.D., Tufts University School of Medicine, Blue Cross Blue Shield of Massachusetts, 401 Park Drive, Boston, MA 02215; e-mail: dana.safran@bcbsma.com. Michael P. Anastario, Ph.D., is with the Cikatelli Associates, Inc., New York, NY. Hector P. Rodriguez, Ph.D., M.P.H., is with the Department of Health Services, UCLA School of Public Health, Los Angeles, CA. Patricia M. Gallagher, Ph.D., is with the Center for Survey Research, Boston, MA. Paul D. Cleary, Ph.D., is with the Department of Epidemiology and Public Health, New Haven, CT. Dale Shaller, M.P.A., is with the Shaller Consulting, Stillwater, MN. William H. Rogers, Ph.D., is with the Tufts University School of Medicine, Boston, MA. Karen Bogen, Ph.D., is with the Mathematica Policy Research, Cambridge, MA.

site-level scores, response rates, the characteristics of respondents, and response patterns between surveys distributed using the different methods.

METHODS

Sample

We selected a sample of 5,648 patients seen by 15 primary care physicians at three care sites in New York. On alternate weeks at each site, we distributed surveys to patients in the office (weeks 1, 3, 6) or mailed them to patients after the visit (weeks 2, 4, 5). Patient age and gender were available for all patients eligible to receive either a handout or a mail questionnaire; other patient characteristics used in analyses were self-reported.

Patient Survey

A version of the Clinician Group CAHPS (C/G CAHPS) survey (Agency for Healthcare Research and Quality, American Institutes for Research, Harvard Medical School, and RAND Corporation 2006) supplemented with items from the Ambulatory Care Experiences Survey (ACES) (Safran et al. 2006) was administered to patients. The instrument included questions about the following: Physician Communication Quality ($k = 6$), Shared Decision Making ($k = 2$), Physical Examination ($k = 2$), Access to Care ($k = 5$), Office Staff ($k = 2$), and Care Coordination ($k = 2$). In addition, the survey included questions about patients' overall rating of the physician and willingness to recommend the physician to family and friends. All survey questions asked about care received over the past 12 months from a particular physician (named in the first survey question).

To generate composite scores for each domain (Physician Communication Quality, Physical Exam, Access to Care, Office Staff, Care Coordination), survey responses for questions in each domain were summed and then the total score was transformed so that the composite scores ranged from 0 to 100 points, with higher scores indicating more favorable performance. If more than half of the items in a composite were missing for a given respondent, the composite score was coded as missing (Nunnally and Bernstein 1994). CAHPS and ACES composites have a physician-level reliability of 0.70 or higher with samples of 45 established patients per physician (Safran et al. 2006; Rodriguez et al. 2007, 2009). The questionnaire was an 8.5 × 11" booklet (8 content pages; $n = 62$ items). Survey packets, which included a cover letter signed by the practice's medical director, the questionnaire, and a postage-paid return envelope, were identical for both distribution methods.

Data Collection

Handout Distribution. Each practice site was instructed to give a survey packet to all patients seen by the participating primary care physicians on weeks 1, 3, and 6 of the study ($n = 2,903$). Patients had the option of returning completed surveys in an on-site drop box or mailing the survey back in a postage-paid business reply envelope.

Mail Distribution. For weeks where the study called for mail distribution (weeks 2, 4, 5; $n = 2,745$) patients were mailed a survey with instructions, cover letter, and postage-paid business reply envelope. Patients who did not return a survey within 2 weeks were mailed a follow-up postcard and a second survey. Mail data collection was conducted by an academic survey research organization, the Center for Survey Research, University of Massachusetts Boston.

Completed surveys were accepted for up to 6 weeks after the second mailing was sent to nonrespondents. A total of 2,591 completed surveys were received. There were 114 duplicate cases where the respondent completed multiple surveys because they made more than one visit during the study period. For these cases, the survey associated with the earliest visit within the study period was used. The final sample included 2,477 completed questionnaires—1,033 from in-office distribution and 1,444 from mail distribution (on average, 165 surveys per physician and 826 per site).

Statistical Analyses

We compared responses in the two distribution modes with regard to distribution rates, cooperation rates, and overall response rates for each week of distribution and total rates. Distribution rates reflect the completeness of survey distribution to the intended sample and were determined by computing the number of surveys administered divided by the number targeted in the starting sample. Surveys undeliverable by mail were excluded from the denominator. Cooperation rates reflect the responsiveness of the population to whom surveys were successfully distributed—calculated as the number of completed surveys divided by the number of surveys distributed. Response rates reflect the rate of completed surveys obtained from the overall target sample—computed as the number of completed surveys divided by the number targeted in the starting sample.

We assessed differences in respondent characteristics by mode. Respondents' age, gender, race, education, self-reported health, and number of visits over an 8-month period were compared by mode using Pearson's χ^2 test for categorical variables, and two-tailed independent sample *t*-tests for continuous variables.

Composite scores were compared between modes using a two-tailed independent samples *t*-test. Ordinary least-squares regression was used to estimate the association between mode and score. Given the clustered nature of the data, we estimated two random effects models where (1) the intercept varied randomly (by physician) and (2) the intercept and slope of the mode effect varied randomly by physician. Top category responses may be particularly prone to mode effects (Elliott et al. 2009). Thus, the percent of respondents choosing the highest response on items and scales were compared by mode using the Pearson's χ^2 test.

To assess whether mail versus handout modes resulted in differential relative standing of physicians and/or care sites, we calculated the correlations of physician relative standing between modes. We estimated a linear regression model that included terms representing the interaction between physician and mode, and site and mode. For each survey measure, a correlation between the intercept and modality slope was estimated using a random effects model which had random effects for the constant in addition to a random effect for the slope (i.e., modality slopes varied by physician). *STATA 9* was used to conduct all statistical analyses (*STATA 9 2006*).

RESULTS

Overall distribution rates were incomplete for handout surveys (74 percent), whereas mail surveys were distributed to the entire selected sample. Distribution rates for handout surveys declined over the 3-week study period and varied by each site (Table 1). Surveys distributed by mail attained a higher total response rate (58 percent; 1,602 completed/2,745 eligible) than handout surveys (40 percent; 1,160 completed/2,903 eligible). Cooperation rates, calculated as the proportion of completes among those who received a questionnaire, were similar for surveys distributed by hand (54 percent; 1,160 completed/2,140 distributed) and mail (58 percent; 1,602 completed/2,745 distributed). Among patients who were handed a survey, only 25 (2.4 percent) returned the survey by drop box versus mailing it in. Lag time between survey distribution and receipt was an average of 8.4 days shorter for surveys

Table 1: Handout Distribution Rate by Office and Week

Completed Surveys	Week			Site Average
	1 (n = 482)	2 (n = 387)	3 (n = 291)	
Site 1	0.63	0.61	0.55	0.60
Site 2	0.86	0.89	0.72	0.82
Site 3	0.85	0.73	0.64	0.74
Weekly average	0.78	0.74	0.64	

Note. Weeks 1, 2, and 3 of the handout mode represent weeks 1, 3, and 6 of survey distribution.

distributed by hand (Table 2). This lag time difference was expected as respondents receiving surveys under the handout mode did not have to wait for their questionnaires to arrive in the mail. Survey respondents were less likely to be under 54 years of age and more likely to be female compared with nonrespondents (Table 2). Respondents for both modes of distribution were older and more likely to be female than nonrespondents, but respondent characteristics did not vary across the modes of distribution (Table 2).

Respondents to surveys distributed by hand reported significantly better experiences on all but three survey items. On average, scores were 2.1 points higher among those who received the survey by handout versus mail (Table 3). In the random intercepts model, there was a significant difference between modes for 19 of the 29 items and scales, with all items comprising the Physician Community Quality composite showing significant mode effects (Table 3). In addition, the percent of respondents choosing the highest response option on items was significantly greater among handout than mail respondents, with particularly large effects observed for physician provision of clear instructions (75.2 percent handout, 66.8 percent mail; $\chi^2 = 16.5, p < .001$) and thoroughness of physical exams (71 percent handout, 64.3 percent mail; $\chi^2 = 11.8, p < .001$) (Table 3).

Composite measures for which physicians' relative standings were most strongly correlated between modes were as follows: Physician Communication Quality ($\rho = .78$), Care Coordination ($\rho = .68$), Access to Care ($\rho = .85$), and Office Staff Quality ($\rho = .76$) (Table 4). Relative standings among physicians for the Shared Decision Making ($\rho = .31$) and Quality of Physical Exam ($\rho = .11$) composite measures were not significantly correlated between modes (Table 4). The most discordant relationship was observed for an item measuring physician provision of comfort during physical exams ($\rho = -.28$).

Several interaction effects between survey mode and individual physicians were statistically significant (Table 4), indicating divergent relative

Table 2: Self-Reported Respondent Characteristics by Mode of Distribution*

	Survey Administration				
	Completed		Distributed		Not Distributed Handout (n = 763)
	Mail (n = 1,444)	Handout (n = 1,033)	Mail (n = 2,745)	Handout (n = 2,140)	
Age*					
18-54	31.9	30.0	40.5	39.0	38.9
55-64	21.1	21.9	16.9	17.6	17.4
65-74	21.3	21.8	15.5	16.0	15.0
75+	25.7	26.3	27.1	27.5	28.7
Gender (% male)*	42.4	39.5	42.9	42.0	42.5
Overall number of primary care office visits					
1	12.1	11.2			
2-4	64.8	65.2			
> 4	23.1	23.6			
Physician-patient relationship duration					
< 1 year	15.6	14.5			
1-3 years	11.1	12.0			
> 3 years	73.3	73.5			
Race (% in group)					
White	98.1	98.3			
Nonwhite	1.9	1.7			
Education (% in group)					
< High school	8.9	9.8			
High school graduate	35.1	35.3			
Some college	31.0	31.9			
College graduate	25.1	23.1			
Days between survey distribution and receipt (SD)	20.4 (10.5)	12.0 (9.4)**			
Self-rated physical health (SD)	53.5 (24.0)	53.7 (23.1)			

Note. Columns presented as percentiles reflect column percents.

* $p < .001$ for comparison between completed and noncompleted surveys.

** $p < .001$ for comparison between mail and handout surveys.

Table 3: Survey Items and Composite Scores by Mode of Survey Distribution

Scales/Items	Unadjusted Mean		Adjusted Mean [†]		Percent Ceiling [‡]		
	Constant (Mail Mode)	Mode Effect	Constant (Mail Mode)	SD (Physician)	Mode Effect	Mail	Handout
Physician Communication Quality	90.7	2.5 ^{***}	90.8	(2.5)	2.2 ^{***}	36.9	41.3*
Explain things	91.6	2.1 ^{***}	91.7	(2.5)	1.9 ^{**}	72.2	77.3 ^{**}
Listens carefully	92.5	2.1 ^{***}	92.7	(2.2)	1.9 ^{**}	76.4	81.3 ^{**}
Clear instructions	89.9	2.8 ^{***}	89.9	(2.4)	2.5 ^{***}	66.8	75.2 ^{***}
Knowledge of medical history	86.7	2.6 ^{**}	86.2	(5.4)	2.0*	60.6	64.9*
MD respects what patient has to say	93.8	2.0 ^{***}	93.8	(2.5)	1.8 ^{***}	80.3	85.3 ^{**}
MD spends enough time with patient	90.5	2.1 ^{***}	90.8	(2.6)	1.8 ^{**}	70.3	75.0*
Shared Decision Making	89.2	2.1*	89.2	(1.9)	2.1*	62.3	66.9
MD gives enough treatment information	8.4	-1.7*	8.4	(1.3)	-1.6	77.8	81.2
MD asked about treatment preference	12.9	-2.5*	13.1	(2.7)	-2.5*	70.2	73.7
Care Coordination	84.2	2.8 ^{**}	83.7	(5.2)	2.5 ^{**}	47.1	50.5
MD talks about medication	89.9	2.6*	89.9	(4.5)	2.2	89.9	92.5*
Follow-up on test results	79.2	2.8*	78.3	(8.7)	2.4	56.8	60.9
Physical Exam	89.9	1.8 ^{**}	89.9	(2.0)	1.5*	58.1	64.4 ^{**}
Thoroughness of MD	88.4	3.0 ^{***}	88.6	(3.8)	2.3 ^{**}	64.3	71.0 ^{**}
MD ensures comfort during exam	90.7	0.5	90.7	(1.2)	0.5	79.9	81.5
Access to Care	78.6	2.8 ^{***}	78.8	(5.8)	2.2 ^{**}	8.1	7.6
Urgent care appointment	83.4	2.4*	83.2	(6.0)	1.5	53.5	56.0

Mail (n = 1,444)

Handout (n = 1,033)

Routine appointment	87.5	2.4 ^{***}	87.6	(2.9)	2.1 [*]	60.6	66.0 [*]
Office calls back	78.2	3.0	78.4	(7.2)	2.4	45.2	46.6
After hours call back	77.8	4.4	77.9	(4.6)	3.9	51.1	56.3
Time spent in waiting room	63.3	3.8 ^{**}	63.3	(13.4)	2.4 [*]	24.8	26.6
Office Staff Quality	87.9	1.9 ^{***}	87.9	(3.6)	1.6 [*]	53.8	56.8
Helpfulness	85.4	2.1 ^{**}	85.4	(4.2)	1.9 [*]	54.1	57.9
Courtesy and respect	90.4	1.6 [*]	90.4	(3.1)	1.3	70.3	73.1
How would you rate your doctor?	89.7	2.3 ^{***}	89.6	(3.5)	2.0 ^{***}	~	~
Would you recommend your doctor?	90.7	2.4 ^{***}	90.8	(3.7)	1.9 ^{**}	~	~

* $p < .05$; ** $p < .01$; *** $p < .001$.

[†]Scores have been adjusted for clustering by MD (random intercepts model).

[‡]“Ceiling” refers to the maximum response option (“Always”) in each response set.

Table 4: Correlations in Physician Relative Standing (i.e., “rank”) between Mail and Handout Modes

<i>Scales/Items</i>	<i>MD-Level Rank Correlations by Mode</i>
<i>Mail (n = 1,444)</i>	
<i>Handout (n = 1,033)</i>	
Physician Communication Quality	0.78***
Explain things	0.76***
Listen carefully	0.66**
Clear instructions	0.81***
Knowledge of medical history	0.81***
MD respects what patient has to say	0.63*
MD spends enough time with patient	0.62*
Shared Decision Making	0.31
MD gives enough treatment information	0.14
MD asked about treatment preference	0.32
Care Coordination	0.68**
MD talks about medication	0.47
Follow-up on test results	0.79***
Physical Exam	0.11
Thoroughness of MD	0.54*
MD ensures comfort during exam	-0.28
Access to Care	0.85***
Urgent care appointment	0.42
Routine appointment	0.70**
Office calls back	0.43
After hours call back	0.14
Time spent in waiting room	0.94***
Office Staff Quality	0.76***
Helpfulness	0.78***
Courtesy and respect	0.74**
How would you rate your doctor?	0.69**
Would you recommend your doctor?	0.59*

* $p < .05$; ** $p < .01$; *** $p < .001$.

standing of physicians across modes. These included the following: the physical exam composite [$F(14, 2210) = 2.2, p < .01$] as well as both of its constituent items—thoroughness of exam [$F(14, 2388) = 1.8, p < .05$] and attention to physical comfort during the physical examination [$F(14, 2210) = 1.9, p < .05$], and the interpersonal item measuring how often the physicians respects what the patient had to say [$F(14, 2394) = 1.7, p < .05$]. For each significant interaction, physicians with lower scores on mail surveys showed substantially elevated scores on handout surveys compared with physicians with higher

performance on mail surveys. The use of 15 physicians limited our ability to thoroughly examine interaction effects, particularly with regard to determining the confidence intervals surrounding correlations between random intercepts and slopes in the mixed effects models. However, only negative correlations were detected between intercepts and slopes. Stable confidence intervals were detected for the thoroughness of physical exam composite measure, where the correlation between the intercept and slope was -0.87 (95 percent CI -0.98 to -0.28).

Finally, an interaction effect was detected at the site level for the Access to Care composite score [$F(2, 2199) = 3.1, p = .04$]. Here, the site with the lowest score in the mail mode showed the largest discrepancy in scores between modes.

DISCUSSION

With patient care experience measures now widely recognized as central to the comprehensive assessment of health care quality, and vital to efforts to advance the goal of patient-centered care, the cost of data collection remains an important rate limiting factor in the uptake of survey-based measures of ambulatory care quality. In-office handout of surveys offers substantial cost savings over mail-based data collection protocols. This study assessing the effects of survey distribution method on survey responses and data quality has several findings relevant to efforts to advance the use of survey-based measurement of patient care experiences.

First, the finding that distribution rates decreased over time for handout surveys suggests process fatigue among office staff during the in-office distribution process. This decline across all sites over time exacerbated the uneven distribution rates by site evidenced at the beginning. These findings suggest that in-office handout distribution of surveys may introduce bias into results. Protocols for mailed survey distribution are more easily standardized.

Second, patient experience scores were significantly higher when surveys were distributed by in-office handout versus mail. Previous research consistently reveals that assessments of experiences vary by modes of survey administration (Walker and Restuccia 1984; Kittleson 1995; Kinnersley et al. 1996; Paolo et al. 2000; Smeeth et al. 2001; Schonlau, Fricker, and Elliott 2002; de Vries et al. 2005; Gribble and Haupt 2005; Rodriguez et al. 2006; Elliott et al. 2009). Modes that increase the proximity of the respondent to the interviewer or visit produce more favorable assessments (Walker and Rest-

uccia 1984; Savage and Armstrong 1990; Kinnersley et al. 1996; de Vries et al. 2005). In our study, surveys distributed by handout had significantly higher scores than those distributed by mail, even after accounting for the clustering of patients within physician practices.

While the correlation of relative physician standing tested for departures from a correlation of 0, correlations in the relative standing of individual physicians varied from -0.28 to 0.94 . With the interactions examined in our study, physicians with lower scores tended to have the most elevated scores in the handout mode. We only had 15 physicians with which to test a physician-mode interaction effect, so it is possible that larger samples of physicians will elucidate such interactions more than we were able to detect in this study. Evidence of interaction effects between physician and mode are particularly troublesome because a simple correction for the “mode” effects associated with handout is not possible under these circumstances.

Our study has some important limitations. First, inference regarding how site-level performance is influenced by setting of survey distribution was limited by only including three sites in the study, which precluded our examining differences in site relative standing as well as site-level reliability. Our ability to detect significant interactions between site and mode of distribution was also reduced, despite the observed interaction for the Access to Care scale. Further mode experiments that incorporate a greater number of physicians and sites may detect and better define the nature of site-mode interactions in this context, and they may produce narrow confidence intervals for the intercept and slope coefficients in random effects models. Finally, results represent those of patients for 15 physicians across three sites and may not generalize to a larger population. Nevertheless, the consistency of distribution mode effects across the survey questions suggests that the findings are robust and are likely to generalize to practice sites structured similarly to the practices involved in the experiment.

In summary, while the cooperation rates and respondent characteristics did not differ significantly by survey distribution mode (handout versus mail), handout rates varied by site and diminished over the 3-week trial in all three practice sites. Surveys distributed by handout mode had significantly higher scores than those distributed by mail, even after adjusting for patient clustering within physicians. We found that the practice site with the lowest mail survey performance had the largest positive handout mode effect and low distribution rates. This suggests that the practice may have been selective in its choice of patients asked to complete the handout survey or that the most favorable patients agreed to take the surveys. Finally, the possibility of physician-mode

or site–mode interaction effects could not be discounted—suggesting that a simple correction for the “mode” effects associated with handout may not be possible. These effects may be exacerbated in a context in which there are “high stakes” associated with the survey scores (e.g., reporting, performance-based financial incentives, etc.). Given the attractiveness of the “handout” mode for many large-scale initiatives, a reasonable next step would be to conduct a test for these mode and physician–mode interaction effects in the context of a “real world” implementation—ideally one in which stakes attached to survey results are high enough to matter to the physicians and/or practices.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This work was funded by the Agency for Healthcare Research and Quality (AHRQ) through the Yale CAHPS Grantee Contract. The authors gratefully acknowledge Angela Li for her generous and dedicated role in project oversight, data management, and manuscript preparation. This work was conducted while Drs. Anastario, Rodriguez, and Safran were employed at the Health Institute, Institute for Clinical Research and Health Policy Studies at Tufts Medical Center and Dr. Bogen was employed at the Center for Survey Research. Dr. Safran (senior and corresponding author) is currently employed as Senior Vice President of Performance Measurement and Improvement, Blue Cross Blue Shield of Massachusetts. Dr. Safran remains an active member of the faculty at Tufts University School of Medicine.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Agency for Healthcare Research and Quality, American Institutes for Research, Harvard Medical School, and RAND Corporation. 2006. “Consumer Assessment of Healthcare Providers and Systems (CAHPS[®]) Clinician and Group Survey: Submission to National Quality Forum.” 2006. [accessed June 10]. Available at <http://www.aqaalliance.org/October24Meeting/PerformanceMeasurement/CAHPSCGtext.doc>
- de Vries, H., M. N. Elliott, K. A. Hepner, S. D. Keller, and R. D. Hays. 2005. “Equivalence of Mail and Telephone Responses to the CAHPS Hospital Survey.” *Health Services Research* 40 (6, part 2): 2120–39.

- Elliott, M. N., A. M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomians, M. K. Beckett, and L. Giordano. 2009. "Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS Hospital Survey Scores." *Health Services Research* 44 (2, part 1): 501–18.
- Gribble, R. K., and L. Haupt. 2005. "Quantitative and Qualitative Differences between Handout and Mailed Patient Satisfaction Surveys." *Medical Care* 43 (3): 276–81.
- Kinnersley, P., N. Stott, T. Peters, I. Harvey, and P. Hackett. 1996. "A Comparison of Methods for Measuring Patient Satisfaction with Consultations in Primary Care." *Family Practice* 13 (1): 41–51.
- Kittleson, M. J. 1995. "An Assessment of the Response Rate via the Postal Service and E-mail." *Health Values* 18: 27–29.
- Nunnally, J., and I. Bernstein. 1994. *Psychometric Theory*. New York: McGraw-Hill.
- Paolo, A. M., G. A. Bonaminio, C. Gibson, T. Partridge, and K. Kallail. 2000. "Response Rate Comparisons of E-Mail- and Mail-Distributed Student Evaluations." *Teaching and Learning in Medicine* 12 (2): 81–4.
- Rodriguez, H. P., T. von Glahn, H. Chang, W. H. Rogers, and D. G. Safran. 2007. "Patient Samples for Measuring Primary Care Physician Performance: Who Should Be Included?" *Medical Care* 45 (10): 989–96.
- . 2009. "Measuring Patients' Experiences with Individual Specialist Physicians and Their Practices." *American Journal of Medical Quality* 24 (1): 35–44.
- Rodriguez, H. P., T. von Glahn, W. H. Rogers, H. Chang, G. Fanjiang, and D. G. Safran. 2006. "Evaluating Patients' Experiences with Individual Physicians: A Randomized Trial of Mail, Internet, and Interactive Voice Response Telephone Administration of Surveys." *Medical Care* 44 (2): 167–74.
- Safran, D. G., M. Karp, K. Coltin, H. Chang, A. Li, J. Ogren, and W. H. Rogers. 2006. "Measuring Patients' Experiences with Individual Primary Care Physicians. Results of a Statewide Demonstration Project." *Journal of General Internal Medicine* 21 (1): 13–21.
- Savage, R., and D. Armstrong. 1990. "Effect of a General Practitioner's Consulting Style on Patients' Satisfaction: A Controlled Study." *British Medical Journal* 301 (6758): 968–70.
- Schonlau, M., R. D. J. Fricker, and M. N. Elliott. 2002. *Conducting Research Surveys via E-Mail and the Web*. Santa Monica, CA: RAND.
- Smeeth, L., A. E. Fletcher, S. Stirling, M. Nunes, E. Breeze, E. Ng, C. J. Bulpitt, and D. Jones. 2001. "Randomised Comparison of Three Methods of Administering a Screening Questionnaire to Elderly People: Findings from the MRC Trial of the Assessment and Management of Older People in the Community." *British Medical Journal* 323 (7326): 1403–7.
- STATA 9. 2006. (*Intercooled for Windows*) [computer program] Version. College Station, TX: Stata Corporation.
- Walker, A. H., and J. D. Restuccia. 1984. "Obtaining Information on Patient Satisfaction with Hospital Care: Mail versus Telephone." *Health Services Research* 19 (3): 291–306.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.