Commentary: Improving response rates to mailed surveys: what do we learn from randomized controlled trials?

Scott D Halpern and David A Asch

Surveying physicians via mailed questionnaires often provides valuable information about physicians’ knowledge, attitudes, and behaviours. Unfortunately, low response rates may introduce bias in that information on those who complete the questionnaire differs from those who do not, in ways related to the outcomes being evaluated.1–3

There is no necessary relation between low response rates and bias.1,2 A 10% response rate in a survey of 100 000 physicians would introduce no bias if the 10 000 responders were similar to the underlying target population in the behaviours or beliefs being evaluated. By contrast, a 90% response rate in this survey might introduce considerable bias if the 10 000 non-responders differed in some important way from the responders. Nevertheless, because higher response rates reduce the potential for bias, and because detecting the bias itself can be challenging, high response rates are viewed by investigators, journal editors, and readers as a proxy for the representativeness of the sample. Higher response rates also lower survey costs per response.1,4 Thus, methods to increase physicians’ response rates to mailed questionnaires are of great interest.

Several interventions to improve response to mailed physician surveys have been tested recently in randomized trials.4–10 In this issue of the International Journal of Epidemiology, Bhandari and colleagues report the first randomized trial to examine whether an endorsement letter from opinion leaders in the same discipline as the respondents improves physicians’ response rate.11 This research question is important for two reasons. First, in our experience, many investigators believe that such endorsement is an effective way to improve response. Such beliefs may be based on observations that opinion leaders are often influential in shaping physician behaviour in other settings, but this effect had not been tested in the setting of mail surveys. Second, if endorsement from opinion leaders was found to work, it would likely represent an inexpensive way to improve response rates.

Bhandari and colleagues tested the effects of endorsement in mailed physician surveys during a study of orthopaedic trauma surgeons’ preferences in the management of tibial shaft fractures. These surgeons were randomly assigned to receive a questionnaire with an endorsement letter signed by 22 respected orthopaedic trauma surgeons, or without such a letter. Surprisingly, the authors found that inclusion of the endorsement letter was not only ineffective, it actually reduced the response rate compared with the group receiving no letter. This conclusion appears to be valid because the study was well-designed, appropriately conducted, and the results were consistent at 2, 4, and 8 weeks. Thus, future investigators must consider what might account for such counterintuitive results, and whether they may be generalizable to surveys in other settings.

First, it may be important to distinguish between Bhandari and colleagues’ intervention—a list of opinion leaders ‘endorsing’ the questionnaire—and an alternative practice of including well-known physicians among the study investigators, and then listing their names on a cover letter to respondents. Although these procedures seem similar, they may be perceived quite differently by respondents. The former practice might be seen as an advertising gimmick, whereas the latter may be seen as providing credibility. Including a respected colleague as an investigator might also invoke norms of reciprocity or otherwise increase the value of responding because doing so might then be construed as doing a favour for a respected colleague.

We suggest these explanations merely as hypotheses—we have no evidence that such distinctions exist. However, previous research has demonstrated that response rates are often sensitive to seemingly minor differences in survey strategy. For example, several studies have demonstrated improvements in response rates based on the type of postage used for the reply envelope: commemorative stamps appear to provide the highest response rate followed, in order, by regular stamps, metered mail, business reply mail, and then no postage at all.10,12 Except for the case where no postage was provided, one might not have expected differences in response rates based on such a trivial choice as postage type.

Second, it is worth pointing out that there is precedence for counterintuitive results in research on interventions to improve physicians’ response rates. One of the first randomized trials in this area found that using outgoing envelopes from a Veterans Affairs hospital, rather than those from a University hospital, produced a higher response rate—a conclusion precisely opposite to the investigators’ hypothesis.5 More recently, we found no effect of including a mint candy with a cash incentive in the outgoing envelope, compared with the cash incentive alone.4 Prior to that finding, we had routinely used mint candies, assuming that the increased bulk they provided would make recipients more likely to open the envelopes and see the cash incentive, rather than throw them away unopened.

1 Center for Clinical Epidemiology and Biostatistics, 2 Center for Bioethics, 3 Center for Education and Research on Therapeutics, 4 Leonard Davis Institute of Health Economics, and 5 Division of General Internal Medicine, University of Pennsylvania, Philadelphia, PA, USA.
6 Center for Health Equity Research and Promotion, Philadelphia Veterans Affairs Medical Center, Philadelphia, PA, USA.

Correspondence: David A Asch, Leonard Davis Institute of Health Economics, 3641 Locust Walk, Philadelphia, PA 19104, USA. E-mail: asch@wharton.upenn.edu

DOI: 10.1093/ije/dyg184
Lastly, it is difficult to estimate the generalizability of this finding among orthopaedic surgeons to surveys of other types of surgeons, let alone to surveys of other types of surgeons, or to studies of physicians in non-surgical disciplines. Indeed, it is even possible that different results would have been obtained if different opinion leaders were chosen from the same field. Perhaps one or more of the endorsing surgeons was a particularly controversial figure, the appearance of whose name decreased colleagues’ willingness to respond.

These speculations are questions of context. Together, they challenge our ability to draw general conclusions about mail survey strategies from a series of separate randomized trials conducted in highly idiosyncratic contexts. Given the sensitivity of response rates to seemingly minor changes in mailing strategy, what appear to be generalizable main effects might really reflect non-generalizable interactions with other contextual elements of the survey.

Is it possible we learn nothing generalizable from these studies when examined individually? Perhaps the only specific conclusion we can draw from any single trial, including Bhandari and colleagues’ trial and those that we and others have conducted previously, is that repeated studies among distinct physician samples are needed to fully explore the influence of any survey design element on response to physician questionnaires. More generally, these surprising findings highlight the fact that even seemingly obvious hypotheses require study, and that such studies ought to involve both two-sided testing and two-sided thinking.13

References