Raising Response Rates: Getting to Yes

Each year, proportionately fewer ordinary people in the United States will give us their time, their records, or their blood. Caught between their enthusiasm and support for biomedical research and their distaste for telemarketing and the other intrusive modern means of selling, people who might have participated in epidemiologic research in previous years now refuse. We sympathize with the growing proportion who just say no, but the challenge of getting adequate response proportions (in the vernacular, "response rates") in the general population demands our attention.

It is tempting to consider the decline too slight to affect interpretation, and in this issue of the journal, Stang *et al*[1] find no marked decline from 1987 to 1995 in Germany. In the United States, it seems that response has been declining steadily for 2 decades; perhaps a longer time period might show a decline in Germany. It also is hard to detect a temporal decline against the noisy background of differences among studies and between men and women, rich and poor, healthy and ill, or retired and employed. One benchmark comes from random digit dialing, a technique that many epidemiologic studies have used as the first stage in sampling individuals. The interviewer chooses a telephone number at random and uses a simple screener to query the individual who answers the telephone (whose name is unknown).

Answering machines, more women working outside the home, and aversion to ever more sophisticated telemarketing have driven down the initial response rates in that setting. More calls are unanswered, and more people hang up before the interviewer gets one sentence out. Tolerable final response usually can be achieved, but only with additional resources to make more calls, to send advance letters to addresses obtained from reverse directories, and to use more skilled and expensive staff to persuade the initially reluctant.

Response to mailed questionnaires, always lower than to other modes of data collection, appears to have fallen only slightly with time. Willingness to be interviewed in person seems to have eroded more, less markedly for respondents who are patients in hospitals or clinic waiting rooms than for respondents at home. Most worrisome is the declining participation proportion for collection of blood or other biologic specimens. Willingness to give blood, which varies enormously with the study population and the setting, appears to have fallen to low levels. In many urban communities in the United States, it takes hard work and good luck to a reach 70% response to a 1-hour in-person questionnaire survey of adults and a 50% response to collection of blood by venipuncture.

Do these response rates threaten validity? The arithmetic of selection bias from non-response is the same as the arithmetic of confounding.[2] If controls with a history of exposure are less likely to participate (a common situation), but cases are about equally likely to respond, regardless of exposure history, bias results. That is, even with similar response in cases and controls, the reasons for non-response typically differ, and the correlations between exposures and non-response easily differ, as well. It is regrettable that many exposures of interest—nutritional habits, use of prescription and over-the-counter medications, sexual and reproductive behaviors—correlate relentlessly with willingness to participate in epidemiologic research.

The new tools of genetic and molecular epidemiology cannot escape the threat of selection bias stemming from non-response. When studying a newly identified polymorphism or serum marker, we are tempted to invoke the protection of ignorance: we know of no reason why the gene or marker would be related to responsiveness. This ignorance is poor assurance of validity. Molecular epidemiology studies with very low response rates will

generate many false leads. How low is very low? In practice, if response exceeds 90% the impact of non-response will be minimal, but response below, say, 50% offers little protection against biases. This is especially true when the two-by-two table contains a small number. Selection bias, like confounding, grows out of associations in the data rather than in the world at large, so a small and wrong number in one cell can reflect a combination of chance and differential reasons for non-response. This situation can arise when most of the subjects decline to give us biological specimens.

If we could measure all of the cultural, class-based, and health-related features that create the correlation of responsiveness and exposure, we could account for these in the analysis and compensate for low response rates, apart from the role of chance. Although it is impossible to measure these features, it is prudent to glean information about non-response from within the study population, or elsewhere, to discern the likely direction and magnitude of the resulting biases.[3]

Better than curing bias from non-response in the analysis is preventing it in the field. Some simple steps can raise response. Looks matter: large, commemorative stamps raise response to a mailed questionnaire,[4] but a professional layout may lower it.[5] A large literature evaluating techniques for improving response to mailed questionnaires offers other tantalizing observations: follow-up telephone calls to non-responders raise the response rate,[6] but telling subjects beforehand that the investigator may call makes them less likely to return the questionnaire unprompted.[7] Furthermore, sending non-respondents a registered reminder letter may work better than a telephone call.[8]

Money helps, too. Cash and other incentives to the study subjects can raise response rates,[9,10] but incentives that are too high risk coercion. Higher pay for the interviewers also may improve the pool of applicants for the critical position of field interviewer.

Often what influences response most is the topic of discussion.[6] In focus groups, respondents often report that a 15-minute segment of the interview went fast if it concerned a matter of interest, for example, their own medical histories, while a 15-minute segment on another topic seemed interminable. This effect of interest in the topic might lead us to conclude that length itself matters little, but that would be wrong. It is easily documented that shorter is better, for response to mailed questionnaires.[7,11] This must be true for in-person or telephone interviews, at least at the extremes, yet research on this aspect of epidemiologic studies is surprisingly thin. The length of the instrument influences the likelihood that the subject will respond at all and the quality of the responses given, two outcomes that are difficult to measure. Review of tape-recorded interviews often reveals that the end of a long interview elicits poor answers that reflect fatigue and irritation, and yet a merely visual inspection of the data easily would deem the interview a "response."

Thus, restraint pays. In part, restraint signifies respect for the participants, as does inclusion of the participants in the products of the research. Participants like to be offered the results of the measurements done on them, such as nutritional analyses of their dietary patterns. They also like the option of receiving a summary of the study findings.

Comfort counts, too. This matters particularly for biological specimens: if the main purpose of collecting a blood sample is to obtain DNA, we may be able to collect buccal cells instead. If blood is required, collection of a very small amount by finger-stick will be met with less resistance than venipuncture. It may hurt just as much, but respondents find it less invasive.

In sum, we find many options available to try to increase response, but they all cost something in information foregone or resources spent. Pilot studies in the months before launching help immensely in finding the approach that leads to good response. Guidance from investigators with similar studies can help even earlier by framing the choices at the point of study design. As we gauge response, do we also need new parameters, beyond those described in Slattery's thorough review[12]? In this issue, Stang *et al* propose a "recruitment efficacy proportion" to be sure we attend to the components of response most within our control. These and other tracking measures may help, but Stang's greater contribution is the synthesis of response data from completed studies.

More research is needed. This seemingly simple recommendation poses problems. Empirical research on epidemiologic methods is not epidemiology, and the underlying phenomena are mutable, culturally mediated human behaviors rather than human biology. The research findings thus tend to be confusing, subtle, and hedged with setting-specific caveats. Furthermore, reports on field methods consume valuable epidemiologic journal space, so they need to be compactly written. But, along with our common sense, they are our best hope for achieving a respectable response rate.

Patricia Hartge
Epidemiology and Biostatistics Program
Division of Cancer Epidemiology and Genetics
6120 Executive Boulevard, EPS 8090
Bethesda, MD 20892

References

1. Stang A, Ahrens W, Jöckel K-H. Control response proportions in population-based case-control studies in Germany. Epidemiol 1999;10:181–183.
2. Rothman KJ, Greenland S. Precision and validity in epidemiologic studies. In: Rothman and Greenland, eds. Modern Epidemiology. 2nd ed. Vol. 1. Philadelphia: Lippincott-Raven, 1998;15–134.
3. Holt VL, Daling JR, Stergachis A, Voight LF, Weiss NS. Results and effect of refusal recontact in a case-control study of ectopic pregnancy. Epidemiology 1991;2:375–379.
4. Choi BCK, Pak AWP, Purdham JT. Effects of mailing strategies on response rate, response time, and cost in a questionnaire study among nurses. Epidemiology 1990;1:72–74.
5. Etter J-F, Perneger TV, Laporte J-D. Unexpected effects of a prior feedback letter and a professional layout on the response rate to a mail survey in Geneva. J Epidemiol Community Health 1998;52:128–129.
6. Wilkins, JR III, Hueston WD, MacCrawford J, Steele LL, Gerken DF. Mixed-mode survey of female veterinarians yields high response rate. Occup Med 1997;47:458–462.
7. Eaker S, Bergstrom R, Bergstrom A, Adami H-A, Nyren O. Response rate

to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. Am J Epidemiol; 1998;147:74–82.
8. Tai SS, Nazareth I, Haines A, Jowett C, A randomized trial of the impact of telephone and recorded delivery reminders on the response rate to research questionnaire. J Public Health Med 1997;1997;2:219–221.
9. Deehan A, Templeton L, Taylor C, Drummond C, Strang J. The effect of cash and other financial inducements on the response rate of general practitioners in a national postal study. Br J Gen Pract 1997;47:87–90.
10. Gilbart E, Kreiger N. Improvement in cumulative response rates following implementation of a financial incentive. Am J Epidemiol 1998;148:97–99.
11. Lund E, Gram IT. Response rate according to title and length of questionnaire. Scand J Soc Med 1998;26:154–160.
12. Slattery ML, Edwards SL, Caan BJ, Kerber RA, Potter JD. Response rates among control subjects in case-control studies. Ann Epidemiol 1995;5:245–249.