

Editorial Comment

JAMIA

When to Use Web-based Surveys

Even in this randomized trial era, surveys of patients and professionals remain an important epidemiologic technique for capturing cross-sectional or longitudinal data, providing fundamental insights about health and disease.¹ There are several alternative methods for collecting such data, which include conducting face-to-face or telephone interviews; circulating a questionnaire by mail, fax, or e-mail; and eliciting responses to a survey posted on an open Web site. Other methods that are currently less widely exploited include digital interactive television; use of a software package to capture survey data, either by mailed floppy disks or as a Java applet over the Web; and automatic telephone menu systems. The preceding article by Schleyer and Forrest² explores some of the issues arising from a Web survey, but it may be useful first to consider the range of survey methods that are available.

The Range of Methods Available, and Their Implications

Two major features define a survey method: how participants are identified and the data capture technique. Let us consider these features and their implications in turn.

How participants are identified determines how much control the investigators exert over their selection. For complete control, eligible participants are selected ac-

ording to specific characteristics.³ We can check characteristics from an existing database or mailing list, including an e-mail discussion list, or carry out a preliminary mailing or interview to capture the necessary eligibility data. If, instead, we wish only a modest degree of control, we can distribute copies of the survey to a group known to us and ask them in turn to send it out to their contacts—the “snowball” technique (e.g., Forsstrom et al.⁴). With this method we have much less control over who receives a copy of the survey, and it makes sending reminders or follow-up surveys more difficult, too. We often do not even know how many copies of the survey have been sent out, making it impossible to calculate a response rate. Thus, generalizing results from a snowball survey is difficult. Finally, there are occasions when we want to encourage anyone and everyone to complete our survey, with no control over the type of participants. We can distribute the survey on the street, in a newspaper, or by random mail shots or we can e-mail it to random members of a general mailing list or place it on a public-access Web site. Such techniques mean that follow-up and reminders are usually impossible and response rates hard to calculate. Schleyer and Forrest used the Web but still retained control over who could respond by issuing passwords by e-mail, allowing selected participants to gain one-off access to a closed Web site. Passwords can also be issued by mail or telephone.

The second key feature is how the data are captured. Many conventional paper questionnaires are designed to capture free-form text, which is usually transcribed. Face-to-face or telephone interviews can be tape-recorded and transcribed, but this is a lengthy process. Such qualitative data can then be analyzed using appropriate software tools such as NUD*IST or MARTIN to

Affiliation of the author: University College, London, United Kingdom.

Correspondence and reprints: Jeremy C. Wyatt, DM, FRCP, University College London, 29 Tabstock Square, London WC1H 9EZ, United Kingdom.

Received for publication: 3/7/00; accepted for publication: 3/21/00.

extract recurrent themes.⁵ Data analysis is easier, faster, and more accurate if data can be captured electronically at source, for example, using specially designed forms and optical character or optical mark recognition, which are increasingly used in clinical trials.⁶ Often the special OCR/OMR forms are faxed to a data center, speeding return of data.⁷ Other techniques for capturing electronic data include telephone voice menu systems, e-mailed text, Web forms, and mailed interactive software packages. As well as speeding data analysis and reducing transcription errors⁸ the choice of data capture method also determines how much control the researchers have over the order of question completion and the time to be allowed. For example, with a paper or e-mailed questionnaire, there is little, if any, control—the participant can leaf back and forth before answering the first question. Even standard HTML forms allow little control over question order or the time allotted, as devious participants can work through the survey answering obligatory questions randomly, then go back to alter their responses. This makes it hard to conduct experiments such as eliciting a diagnosis on a simulated case before and after receiving the output of a decision support system.⁹ On the other hand, face-to-face or telephone interviews, interactive software, and Java applets delivered via the Web allow complete control over the time participants are allowed and the ordering of questions.

The combination of these two features determines three further significant issues:

- The type of data that can be collected: responses only (postal, e-mail, Web—because of variable delays), responses and time taken to respond (automatic phone system, software, recorded interview), responses and observer comments (face-to-face encounters like a conventional job interview)
- The overall survey cost: high (face-to-face, interactive software), medium (postal surveys), low (e-mail, automatic phone system, Web)
- The speed of data collection and time to completion: fast (Web, e-mail), medium (fax, phone, face-to-face), slow (postal surveys)

With this general background in mind, let us consider the benefits and problems posed by Web-based surveys of the type described by Schleyer and Forrest.

Benefits of Web-based Surveys

Web-based surveys have a number of benefits over conventional paper or face-to-face methods. These include the following:

- They are more inclusive, allowing a further reach than postal or phone surveys or direct interviews, potentially including a global audience. This can be useful in finding reasonable numbers of respondents with a rare condition.
- Once set up, they are cheap to carry out, making it easier to recruit large numbers of participants or to collect data repeatedly, on several occasions.
- The data are captured directly in electronic format, making analysis faster and cheaper. This again allows more data to be collected than with conventional mailed paper questionnaires.
- Associated material, such as data definitions or even the protocol for the study, can be linked to the data capture forms and vice versa.¹⁰
- They allow interactive data capture with rapid checking of responses, at least at the form level; immediate validity checking of individual data items requires a Java applet.
- Web surveys allow the use of multimedia and enforced branching, and with Java applets they allow complex experiments with complete control over the scheduling of stimuli and responses without the need to mail each participant a floppy disk. However, there is a problem with measuring timing using simple HTML forms, as network response times are highly variable.¹¹
- Web surveys allow rapid updating of questionnaire content and question ordering according to user responses. This can be useful, for example, in Delphi studies.

Problems with Web-based Surveys

The two key disadvantages of Web-based surveys concern the generality and validity of their results. The generality of the results is clearly restricted to those who are keyboard and Internet literate—currently only a third of the population. Second, while it may be easy for *JAMIA* readers to understand what is required of them in a Web survey, it is not for everyone—the paper by Schleyer and Forrest has some sobering examples. A third problem is that, because of simple preference or shortage of time in the office, some participants will prefer to print off the survey document to complete on the train, on a plane, or at home. Unless this is allowed, such participants will be excluded from the group, potentially biasing the results. However, perhaps one of the most worrying threats is that a keen participant can respond multiple times to a survey, shifting the average results in their favor, or may even recruit their friends by sharing

passwords. An early Web survey conducted by Byte, of the proportion of operating systems installed on desktop machines, makes an instructive example.¹² The results showed that AIX was apparently the most popular operating system. However, when the surprised Byte staff reviewed the server log, they discovered that not only was IBM.COM the most frequent domain but that many participants had answered the survey on several occasions, the record being 80 times! This problem can be eliminated by issuing participants with passwords that are valid only once—as in the article by Schleyer and Forrest—but this takes software expertise and requires administration. A further threat to generality is that some participants may be reluctant to complete the survey unless there is a guaranteed anonymity. Thus, in future, we may need to provide the same secure Web servers and encryption for Web surveys that we provide for patient record systems,¹³ especially if participants can be identified (which they often can¹⁴) and the survey examines sensitive issues.

Although a response rate above 80 percent is usually vital to ensure the generality of survey results, it is not always necessary. For example, in the unusual circumstances that the community surveyed are homogeneous with respect to a key variable,¹⁵ or when documentation of the simple occurrence of a bizarre or rare phenomenon is of interest (such as three surgeons who are heroin addicts), a lower response rate is less of a problem.

However, the major concern is whether the data captured using Web surveys are reliable and valid. Reliability means that the same question should elicit the same answer on two occasions from the same person. Validity means that the question is measuring what it is intended to measure³ and contributes distinct information to a scale measuring a complex concept.¹⁶ Establishing the validity of a question or a scale derived from several questions requires many careful studies; this explains why it has only recently become feasible to measure intuitive concepts such as quality of life using paper questionnaires.¹⁷ Unfortunately, simply translating the format from paper to the Web may lead to significant changes in the perception of what the questions and answers mean¹⁸ and, thus, the validity of the survey. The Schleyer and Forrest paper gives examples of problems affecting validity, including coding errors and problems experienced by AOL subscribers. The investigators were unable, however, to exclude the possibility that other ISP users might have had more subtle but equally important problems affecting their responses. Simple errors that may reduce data validity are more likely in Web than paper surveys, including participants' not scrolling down to

see a whole page of questions or list of options in a list box and not understanding how to correct a mistaken response. These are more acute with the Web than with paper questionnaires, because each user experiences a subtly different questionnaire according to their screen size, hardware platform, operating system, browser, and Internet service provider. Participants may even have changed their default screen colors or fonts in a way that obscures significant detail in the questionnaire. The fact that people select responses with a mouse rather than a pen means that existing paper survey instruments, such as those for quality of life, may need to be revalidated on screen. Equally, the ability to copy and paste response text from anything on their personal computer or the wider Web may perhaps lead to new patterns of response that fail to reflect true feelings. One potential solution to many of these difficulties would be to expose all participants to a screen of "calibration" questions before the survey proper. If their responses to these are satisfactory, they would be allowed access to the full survey; if not, their responses would be carefully scrutinized.

A further difficulty with Web surveys is that they may be harder to validate than questionnaires conducted face to face or with local participants.¹⁹ Piloting will be just as important as with paper surveys,²⁰ so researchers should stay at the pilot participants' side while they use the form on screen, to listen to comments, detect and log their misconceptions, and attempt to correct them. Such piloting was carried out in the survey by Schleyer and Forrest, but it is still worrying that a significant software error was discovered only after 130 sets of data had been collected.

One problem with Web surveys is that it may be more difficult to carry out repeated assessments of the same individual for epidemiologic purposes. In a study of patients with ulcerative colitis,²⁰ 70 percent claimed to have had the same e-mail address for two years, but this is a very short period in epidemiological terms.

A chronic problem in surveys is poor response rates, perhaps due to survey fatigue or even reaction against "survey serfdom."²² This may also affect Web surveys.²⁰ Response rates usually matter a great deal, and a rate above 80 percent is usually necessary to ensure validity (and publication in the better journals). Reminders are a key way to enhance response,¹⁹ but there is always a concern that the quality of data returned in the second and third rounds may have been reduced by the irritation of participants on receiving paper or e-mail reminders. Schleyer and Forrest found that their group needed a lot of pushing to fill out the Web form, and sent over 1,100 reminder e-mails; this

potentially represents a considerable amount of effort. With paper surveys, another important technique to enhance response is incentives, such as a draw for prizes or small cash amounts.²³ With a probable international range of recipients, it is unclear how to provide such financial incentives in Web surveys. Eysenbach and Diepgen²⁴ provided eczema patients with an incentive to complete a Web survey by offering them an eczema score on completion, but we do not know how effective such incentives are.

Conclusions

It is clear that we must be cautious about introducing further uncertainty into our current research methods, which in medical informatics are already so poorly validated.³ There is currently a risk that, if widely adopted in medical informatics, Web surveys could threaten the generality of some studies and distance us from the silent majority of patients and clinicians who are still more familiar with reading and recording data on paper than on screen. The choice about whether to use the Web for a survey should thus not be driven by economics but by a consideration of the many alternative techniques and other issues discussed here. However, when we do decide in the future that a Web approach is unsatisfactory, we may need to justify our position to funding agencies, always keen to economize.—JEREMY C. WYATT, DM, FRCP

The author thanks Patti Brennan, University of Wisconsin, for her support for this article.

References ■

- Sackett DL, Wennberg JE. Choosing the best research design for each question. *BMJ*. 1997;315:1636.
- Schleyer TKL, Forrest JL. Methods for the design and administration of Web-based surveys. *J Am Med Inform Assoc*. 2000;7:416–25.
- Friedman C, Wyatt J. *Evaluation Methods in Medical Informatics*. New York: Springer-Verlag, 1997.
- Forsstrom J, Rigby M, Roberts R, et al. Key Recommendations of the TEAC-H Project: Toward Evaluation and Certification of Telematics Services for Health. Turku, Finland: University of Turku, 1999.
- Tesch R. Computer programs that assist in the analysis of qualitative data: an overview. *Qualitative Health Res*. 1991;1:309–25.
- Hosking JD, Newhouse MM, Bagniewska A, Hawkins BS. Data collection and transcription. *Controlled Clin Trials*. 1995;16:66S–103S.
- Taylor DW, Bosch EG. The DataFax project [abstract]. *Controlled Clin Trials*. 1991;11:296.
- Wyatt JC. Acquisition and use of clinical data for audit and research. *J Eval Clin Pract*. 1995;1:15–27.
- Friedman CP, Elstein AS, Wolf FM, et al. Enhancement of clinicians' diagnostic reasoning by computer-based consultation: multisite study of 2 systems. *JAMA*. 1999;282:1851–6.
- Afrin LB, Kuppaswamy V, Slater B, Stuart RK. Electronic clinical trial protocol distribution via the WWW: a prototype for reducing costs and errors, improving accrual and saving stress. *J Am Med Inform Assoc*. 1997;4(1):25–35.
- Wood FB, Cid VH, Siegel ER. Evaluating Internet end-to-end performance: overview of test methodology and results. *J Am Med Inform Assoc*. 1998;5:528–45.
- BYTE learns a lesson in Web surveys. *Byte*. Sept 1996:32.
- Rind DM, Kohane IS, Szolovits P, Safran C, Chueh HC, Barnett GO. Maintaining the confidentiality of medical records shared over the Internet and the World Wide Web. *Ann Intern Med*. 1997;127:138–41.
- Frisse ME. What is the Internet learning about you while you are learning about the Internet? *Acad Med*. 1996;71:1064–7.
- Leslie LL. Are high response rates essential to valid surveys? *Soc Sci Res*. 1972;1:323–34.
- Coste J, Fermanian J, Venot A. Methodological and statistical problems in the construction of composite measurement scales: a survey of six medical and epidemiological journals. *Stat Med*. 1995;14:331–45.
- Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF 36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ*. 1993;306:1440–4.
- Wyatt JC. Same information, different decisions: format counts [editorial]. *BMJ*. 1999;318:1501–2.
- Oppenheim AN. *Questionnaire design, interviewing and attitude measurement*. London, UK: Pinter Publishers, 1991.
- Suchard MA, Adamson S, Kennedy S. Piloting patient attitudinal surveys on the Web. *BMJ*. 1997;315:529.
- Soetikno RM, Provenzale D, Lenert LA. Studying ulcerative colitis over the World Wide Web. *Am J Gastroenterol*. 1997;92:457–60.
- Phillips M. Cease survey serfdom. *Lancet*. 1994;343:1236.
- Deehan A, Templeton L, 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.
- Eysenbach G, Diepgen TL. Epidemiological data can be gathered with the World Wide Web [letter]. *BMJ*. 1998;316:72.

APPENDIX

*Draft Guidelines for Web-based Surveys***Scenarios that may be suitable for a Web-based survey:**

Respondent features:

- Respondents are already avid Internet users; e-mail addresses known for reminder messages.
- Respondents are enthusiastic form fillers, will not require monetary incentives.
- Need for respondents covering a wide geographic area (e.g., rare clinical specialties, diseases)
- Respondents known to match nonrespondents and even non-Internet users on key variables.

Survey features:

- Need for complex branching, interactive questionnaire, or multimedia as part of the survey instrument.
- Survey content will evolve fast (e.g., Delphi surveys).
- Intent is to document bizarre, rare phenomena whose simple occurrence is of interest.
- No need for representative results: collecting ideas vs. hypothesis testing.

Investigator features:

- Limited budget for mailing and data processing but good in-house Web skills.
- Precautions can be taken against multiple responses by same individual, password sharing.
- Web survey forms have been piloted with representative participants and demonstrate acceptable validity

and reliability with most platform/browser/ISP combinations.

- Data are required fast in a readily analyzed form.

Scenarios unsuitable for a Web-based survey:

Respondent features:

- Target group is under-represented on Internet, e.g., underprivileged or elderly people.
- Target group is concerned, however unreasonably, about privacy aspects.
- Target group requires substantial incentives to complete the survey.
- Need for a representative sample.

Survey features:

- Need for accurate timing or observational data on participants.
- An existing paper instrument has been carefully validated on target group.
- Need to capture qualitative data or observations about participants.
- Need to capture accurate timings (unless Java applets used).
- Wish to reach the same group of participants in the same way months or years later.

Investigator features:

- Limited in-house Web or Java expertise, but existing desk top publishing and mailing facility.

**ERRATUM****Error in Wording**

On p. 159 of the Mar/Apr issue of *JAMIA*,¹ in the last line of the hypothetical Dextractor query sequence in the left-hand column, the word OR should replace the word AND, so that the full line reads:

Set 8: Boolean: 6 OR 7; 342 patients

1. Nigrin DJ, Kohane IS. Temporal Expressiveness in Querying a Time-stamp-based Clinical Database. *J Am Med Inform Assoc.* 2000;7(2):152-63.