Advances in Nursing Science
Vol. 26, No. 4, pp. 246–256

Measurement Issues Related to Data Collection on the World Wide Web

Ora L. Strickland, RN, PhD, FAAN; Margaret F. Moloney, RN, PhD, ANP; Alexa S. Dietrich; Stuart Myerburg, JD; George A. Cotsonis, MA; Robert V. Johnson

As the World Wide Web has become more prominent as a mode of communication, it has opened up new possibilities for research data collection. This article identifies measurement issues that occur with Internet data collection that are relevant to qualitative and quantitative research approaches as they occurred in a triangulated Internet study of perimenopausal women with migraine headaches. Issues associated with quantitative data collection over the Internet include (a) selecting and designing Internet data collection protocols that adequately address study aims while also taking advantage of the Internet, (b) ensuring the reliability and validity of Internet data collected, (c) adapting quantitative paper-and-pencil data collection protocols for the Internet, (d) making Internet data collection practical for respondents and researchers, and (e) ensuring the quality of quantitative data collected. Qualitative data collection over the Internet needs to remain true to the philosophical stance of the qualitative approach selected. Researcher expertise in qualitative data collection must be combined with expertise in computer technology and information services if data are to be of ultimate quality. The advantages and limitations of collecting qualitative data in real time or at a later time are explored, as well as approaches to enhance qualitative data collection over the Internet. It was concluded that like any research approach or method, Internet data collection requires considerable creativity, expertise, and planning to take advantage of the technology for the collection of reliable and valid research data. **Key words:** *data collection*, *internet*, *measurement*, *questionnaires*, *World Wide Web*

THE ESTABLISHMENT of the Internet, and specifically the World Wide Web, as a common mode of communication during the past 2 decades has greatly enhanced interaction among individuals and groups in all areas of life. Although the business community readily recognized and integrated Internet communication into its daily mode of operating, the nursing and research communities have been more cautious in their use of this communication tool in their activities. The major use of the Internet in nursing has been as a resource locator or for data collection with demographic surveys.[1] While researchers have found the Internet to be quite useful for gaining access to information by searching the World Wide Web, they have been more reluctant to use it to conduct one of its core activities—collection of research data from respondents. Undoubtedly, Internet data collection can have an important role in research. Data collected over the Internet are likely to be less affected by social desirability and inhibition than that collected via paper-and-pencil methods.[2] Although the World Wide

From the Nell Hodgson Woodruff School of Nursing (Drs Strickland and Moloney), Department of Anthropology (Ms Dietrich), and Rollins School of Public Health (Mr Myerburg, Cotsonis, and Johnson), Emory University, Atlanta, Ga.

This article is based on a study supported by NIH Grant No. R15 NR05303.

Corresponding author: Ora L. Strickland, RN, PhD, FAAN, Nell Hodgson Woodruff School of Nursing, 1520 Clifton Rd. N.E., Atlanta, GA 30322 (e-mail: ostric@emory.edu).

Web can offer researchers global access, fast interaction with subjects, and automation,[3] the facility and adeptness with which data are collected will depend upon the type of data collected (ie, qualitative vs quantitative), the computer skills of the respondent, and the computer hardware and software available.

The primary goal of measurement of any phenomenon is that the data collected are consistent, accurate, and as precise as possible. It is crucial to adequately address these goals no matter whether qualitative or quantitative data are obtained. However, when data are collected via the Internet these aims can be uniquely addressed and challenged regardless of the type of data obtained. This article will address measurement issues related to both qualitative and quantitative research approaches, as they occurred in a triangulated Internet study of perimenopausal women with migraine headaches.

AN INTERNET STUDY OF MIGRAINES IN PERIMENOPAUSAL WOMEN

The purpose of the study was primarily to investigate the experiences of perimenopausal women with migraine headaches, the quality of life experienced by this group, and the self-care interventions used for their headaches. A secondary purpose was to examine the feasibility and effectiveness of collecting data via the Internet. Women were recruited who met criteria for migraine headaches and perimenopause, and who were willing to participate in the study using the Internet. The study used mixed methodology, including a quantitative health history tool, a migraine-specific quality-of-life tool (MSQOL),[4] and SF-36,[5] all of which were completed on-line, and an on-line discussion board, which was to function as a focus group. In addition, real-time qualitative interviews were performed to ensure that adequate data were obtained to meet the primary study purpose.

It was initially planned that participants would be recruited from a university community where the study was located, which included faculty, staff, and students who had e-mail access to the Web site. This was done using flyers, e-mail contacts, and word of mouth. Because of difficulties in recruitment, the population was expanded to include individuals in the larger community, including those accessible by Internet. Links were placed from the study Web site to university search engines, and later to an menopause Web site. Participants volunteered for the study by e-mail or phone. Several individuals volunteered by downloading the sample consent form from the Web site, completing it, and mailing it to the project investigator. Initially, the consent form was completed before the participant was screened. In order to simplify the process, this was later modified so that individuals were screened on the phone before completing and mailing their consent forms.

The research project was a collaborative endeavor of nurse researchers and specialists in information services. Although the researchers had access to the Web sites, the information services personnel were responsible for making changes in the web page content or format on the basis of the suggestions of the researchers.

Following screening and completion of the consent form, each participant was given a login name and password for the questionnaire section on the Web site. She also received a folder containing a card with her login name and password, and information about accessing the Web site, completing the questionnaires, and participating in the discussion board. The questionnaires were then completed at the convenience of the participant during one or multiple Internet sessions. At the completion of the final questionnaire, a programmed notification was sent by e-mail to the researchers, informing them that the participant had completed the questionnaires.

When 4 to 8 participants had completed their questionnaires, a new discussion board was created and placed on the Web site, along with a new login name and password, which was then given to the participants for

this group. They were also asked to choose a pseudonym for the discussion board, and were again given instructions to access and use the board. Discussion boards were initially planned to be active for 3 weeks; however, when participation was slow, they were allowed to remain active for a longer period. One discussion board was active for about 5 weeks.

Upon completion of the discussion board, each participant was contacted by telephone and asked to complete a short telephonic interview regarding her perceptions of participation in an Internet study. This included perceptions regarding ease of completing the questionnaires and the discussion board, and identification of any problems or suggestions about the process. Participants were paid a small incentive fee at the completion of the evaluation interview.

ISSUES IN QUANTITATIVE DATA COLLECTION

Quantitative data collection over the Internet can involve the collection of survey data or data for empirical studies,[6] such as for psychological studies or for health assessments. Use of questionnaires is a common method of data collection over the Internet since they tend to consist of structured and closed questions. Questionnaires are particularly convenient for Internet use because they can be easily posted to newsgroups, e-mail lists, e-mailed directly to participants,[7] or placed on a Web site or gopher site for download.[8] When collecting quantitative data over the Internet, existing measurement instruments can be adapted for use over the World Wide Web. However, survey questionnaires and the data collection protocol selected need to be (a) appropriate for addressing the aims or purposes of the data collection, (b) reliable and valid, (c) adaptable for use over the Internet, and (d) practical for respondents and the researcher. In addition, consideration needs to be given to ensuring the quality of data collected over the Internet. In this study, existing questionnaires were adapted for loading onto the Internet. One instrument, the SF-36,[5] has been widely used as a paper-and-pencil instrument, but is also now increasingly used on the Internet. The other instruments included a descriptive health history that was adapted for this study from a previous study,[9] and the MSQOL.[4] Both were modified in format for use on the Internet.

Data collection aims

When data are to be collected for research over the Internet, the research purposes and aims, including research questions and/or hypotheses, specify the data that will be required and the target population as is the case with any research investigation. However, prior to initiating data collection over the Internet, a complete and thorough protocol for data collection needs to be specified, that integrates the steps required to successfully place each survey questionnaire on the Internet and to retrieve data from the Internet. Each variable for which data are required should be linked to a data collection form or questionnaire that is appropriate for the target population. The Internet data collection protocol needs to (a) provide the details of how potential participants will be screened to ensure that Internet data collection is appropriate, (b) specify how respondents will be oriented and gain access to the data collection Web site, (c) ensure that respondent identification numbers are linked to each data source, (d) specify how respondent confidentiality will be maintained over the Internet, (e) indicate when and in what order each questionnaire should be completed, (f) note how and in what form Internet data will be retrieved for analysis, and (g) specify how Internet data will be permanently removed from the Web site. A major advantage of Internet data collection is that questionnaires and forms can be set up on the Internet so that data can be downloaded directly into data management programs. The development of an Internet data collection protocol will require the involvement of an information

services specialist who is informed about the research protocol and who has experience with setting up Web sites for data collection that can be easily entered, downloaded, and kept secure and confidential.

For the migraine study, the study protocol provided a number of steps to ensure that respondent confidentiality would be maintained, questionnaires would be successfully completed, respondents would be capable of completing questionnaires over the Internet, and data would be readily available for analysis. Initially, when potential participants were screened, they were questioned about their access to computers and their ability to do the computer work required to provide data. Participants who had limited experience with computers and the Internet were provided assistance either in the form of an undergraduate student assistant, or via telephone and/or e-mail assistance from the researchers. Respondents were also given study packets that included sample screen Web site pages for the study, with instructions and passwords. When a participant was admitted into the study, her identifying information was entered into a computerized database that then provided a login name and password based on her personal data. This database, which was connected to the questionnaire site, permitted access for the respondent to complete the questionnaires on-line. Each questionnaire was then accessed. When the questionnaires were completed, the data were then automatically transferred to an Excel file, which could be imported into SPSS for analysis. Firewalls and secure socket layers were used to maintain confidentiality of the databases. An individual study number identified participants in the analysis software databases.

Reliability and validity issues

Questionnaires and forms that are to be placed on a Web site for data collection are only as good as their prior evidence for reliability and validity. When selecting instruments for Internet data collection, it is best to start by selecting those that have prior evidence that they have metric properties of high quality. A concern that the investigator must consider is whether the Internet is a good forum for collecting the data for each measure and if this approach is likely to influence the quality of the data obtained. A reality of measurement that should not be forgotten is that instruments that have been reliable and valid in one situation may not function well under other circumstances. In most cases, questionnaires that have been shown to be reliable and valid for self-report using paper-and-pencil methods are also likely to be reliable and valid for completion over the Internet. However, one should not assume this is the case. Ideally, the reliability and validity of questionnaires administered over the Internet should be assessed prior to using them in a major Internet study. At the very least, the reliability of questionnaires with formats amenable to internal consistency reliability assessment needs to be calculated once data have been collected. Lack of internal consistency would be an indication that the data are either not reliable or lack insufficient variability. Questionnaires that are not reliable are also not valid.

Questionnaires and forms that have complicated or confusing formats may present special problems for the webmaster in placing these on an Internet site and are also more likely to be difficult for respondents to complete accurately. Reliable and valid quantitative data collection over the Internet requires that respondents have the ability to navigate the Internet to the extent that they can access the Web site. In addition, they need to have the ability to read and enter data at a level appropriate for the reading and comprehension levels of questionnaires placed on the Web site. In our study, the questionnaires underwent a test period before the actual data collection began. Undergraduate students participated in a trial use of the Internet site, which included completion of the questionnaires. The students then provided feedback about the challenges and benefits of completing the questionnaires over the Internet. The health history, a relatively straightforward

questionnaire that included history of other illnesses and medication use, contained a series of questions with possible answers, some open-ended questions, and some questions that required making lists. This descriptive questionnaire had been used without difficulty in a previous study as a paper-and-pencil version.[9] The practice sessions revealed that the format developed for entering medications was somewhat confusing when completed over the Internet, and this was revised. The SF-36 and MSQOL both consisted of simple Likert-type scales, and were easily formatted in a way that resembled the paper-and-pencil versions. A problem that arose later in actual data collection was that participants who had older web browsers were sometimes unable to advance through the questionnaire from page to page. The format was then revised to accommodate this problem, and a note was placed on the questionnaire site that recommended the use of newer browsers, if possible.

Other issues that may affect the validity of Internet data are data fraud, data generalizability, and self-selection bias.[10] It is possible that in some studies, individuals who complete questionnaires over the Internet could complete them multiple times. To deal with this situation, software can be set up to identify the domain address so that multiple sets of responses from the same web address can be identified. For the migraine study, each questionnaire when completed could not be repeated by that respondent because identifying information was entered by the researchers that informed them when the data collection was completed and subjects were then prohibited from duplicate completions. No one was admitted into the questionnaire site on the Internet without receiving her individualized login name and password.

The issues of data generalizability and selection bias are related. It is likely that those who participate in an Internet-based study are different from the general population because those who cannot afford a computer are less likely to participate. The extent to which members of the target population for a study has computer access could result in selection bias by making the sample more affluent than the population of interest. In this study, many of those recruited had found the study on the Internet, meaning that they already had computer expertise. Other participants were recruited from flyers and word-of-mouth in the local community. Several of these participants had little computer experience but with support were able to access the Web site and complete the instruments. One participant had very little computer expertise and did not have direct access to a computer. An undergraduate assistant met with this participant in a local public library, helped her create a free e-mail account, and assisted her in logging on to the Internet to complete the questionnaires and the discussion board. Therefore, public libraries with Internet access is a viable option for including respondents who cannot afford a computer in an Internet-based study, thereby reducing selection bias.

Adapting quantitative data collection protocols for the Internet

As noted above, existing questionnaires are often used for quantitative data collection on the World Wide Web. Support for interactive documents, or "forms," has been made possible through the use of HTML, which made it possible for respondents to complete questionnaires over the Internet. With the development of HTML, it became possible to place questionnaires on the Internet that are visually and functionally identical to conventional questionnaires. Individuals completing questionnaires over the World Wide Web can check questionnaire boxes, radio buttons, and text-entry boxes with ease and minimal chance of error.[11] For most forms and questionnaires, directions to respondents will need to be revised to accommodate the Internet. For example, provisions need to be made for respondents to easily check and change their responses if they desire, prompts that encourage completion of all items can be integrated (unless intentionally left blank by the respondent), and special links between

forms may be required. Questionnaires can be placed on the Internet in a manner to reduce errors and missing data, such as analyzing responses prior to final acceptance to eliminate errors and inconsistencies in responses,[11] embedding answer reliability checks in the questions, and using forced-choice answers for questionnaires or rating scales[12] along with an option that notes that an item was intentionally unanswered. Data collected over the Internet have been shown to have less missing data when questionnaires have been carefully designed for Internet data collection.[13]

If a longitudinal study is being conducted, the timing of data collections and communication of data collection points for the participants are particularly important. The investigator will need to set up an approach to remind study participants when to complete follow-up data collections at the scheduled times. In addition to reminders for timed data collections, it may be necessary to remind participants to complete data collection when they are completing questionnaires at their leisure. This can be facilitated through e-mail reminders that are linked to the Web site. It may be necessary to get post office addresses in addition to e-mail addresses since some participants may not have access to their own personal computer.

In our study, we found that participants often needed to be reminded to complete questionnaires and log on to the discussion board. Weekly (or twice-weekly) phone calls were made to remind participants to complete the questionnaires and to ask if they needed assistance. Also, e-mails were used for this purpose, and "e-cards" were sent occasionally, as a way to tactfully remind participants that they needed to complete the work.

The practicality of Internet data collection

The best data collection protocols are practical for subjects as well as for the investigator. Instruments for quantitative data collection are practical for respondents when they are accessible, appropriate for the target population, easy to understand and complete, simple, and not demanding of energy and time.[14] This is also true for collection of data over the Internet. When deciding to use the Internet for data collection, the investigator must carefully consider the characteristics of the target population. As noted previously, the literacy of the target population needs to be at a level where completion of questionnaires over the Internet would not be a problem. In general, questionnaires and forms should not be above the fifth grade reading level. Another characteristic that is important is the health of the target population. Do potential respondents have mental or physical limitations that are likely to compromise Internet data collection? Respondents who are young children, ill, or frail may find Internet data collection too challenging, although Fleitas[15] found that chronically ill children were actively involved participants. The investigator needs to consider the amount of time required for completion of data collection over the Internet as well. When too much energy and time are required of respondents to complete data collection, the dropout rate will increase.[16] An advantage of Internet data collection is that it is possible to set up the Web site so that respondents can return to the Web site and complete questionnaires at their leisure in case they become fatigued. If this is done, a mechanism to remind respondents to return to the Web site to complete questionnaires will be required so that the amount of missing data would not increase.

Although many individuals have personal computers at home or at their place of employment, access to a computer may be a practical barrier to participation for some potential respondents. As noted above, access to the Internet is often available through public libraries; however, this requires participants to be willing to travel to another site to participate in the study.

Measurement protocols and instruments are practical for researchers when they are accessible, easy to administer and score, and not too demanding of researcher time and other resources. When quantitative data are

collected over the Internet it is important to remember that the associated costs not only include the computer equipment and software, but also the technical expertise of an information services specialist and support personnel to assist with monitoring the Web site. If the Internet survey uses a standardized instrument that is usually purchased, the cost of getting approval to place it on the Web site must be considered, as well as the cost of scoring. Quantitative data collection over the Internet may be more costly in regards to purchasing equipment and computer expertise and setting up the Web site, but less costly in regards to the expense of recruiting participants, contacting participants, and data management, particularly when large samples are needed.[3,17,18] Since the Internet allows for reduced cost, global access, and real-time data collection,[3] the potential size of clinical and survey studies that can be practically implemented is greatly increased because the human effort required and the financial cost of data collected over the Internet are greatly reduced.

Ensuring the quality of quantitative Internet data

A study's quality is indelibly linked to the quality of the data collected. When study participants provide data over the Internet, they need to be adept at using the computer. It is not unusual to have individuals decide to participate in an Internet study who do not have adequate computer skills to navigate the Internet to get to the Web site. They also may not know how to respond to questionnaires and forms on the Web site. Therefore, training of respondents is necessary, and may need to be done in real time and space before the individual attempts to access the Web site. When clinical data are collected over the Internet, clinicians need to be trained to directly enter data into the Internet database to increase data quality. Real-time data entry by clinicians can reduce transcription errors and "semantic errors as the clinicians managing the patients are directly entering data themselves."[3]

Maintaining the quality of research data is difficult, particularly when large clinical data sets are being compiled. Data must be carefully and continuously recorded, collated, entered into a database, and validated.[3] Data collected over the Internet can facilitate audit and progress charting to ensure the quality and completeness of study data. Descriptive statistical reports based on Web site data can be more efficiently and effectively generated on a regular basis throughout the study period to monitor study progress. This should include recording the number of respondents, percentage of respondents completing questionnaires at each data collection point, and the amount and nature of missing data. When Internet questionnaires are used, there is less risk of coding errors and data can be easily downloaded for use with other software for validation, reliability checks, and analysis. Other strategies for ensuring data quality have already been mentioned. These include (a) integration of Web site prompts to encourage participants to complete all questions and questionnaires, (b) review of the Web site frequently to ensure that participants have completed questionnaires at the proper time, and (c) assessing instruments with scales for internal consistency reliability using data generated within the study.

ISSUES IN QUALITATIVE DATA COLLECTION

When qualitative data are collected over the Internet, it is imperative that the process adheres to the philosophy of the approach used. Prior to selecting the Internet as a means for data collection, the researcher should consider whether Internet data collection allows the research process to remain true to the philosophy of the qualitative approach desired. It is possible with some approaches that all the data required for the study could be collected over the Internet. This may be the case with many phenomenological studies that focus on the meaning of one's lived experience[19]; interviews with some individual participants who are able to

articulate their thoughts in writing may be appropriate in some studies. Qualitative focus groups may also be appropriate, as we discuss below. Other approaches may only be amenable to having part of the required data collected over the Internet. For example, part of the data collected for a grounded theory approach could be obtained via the Internet because this approach focuses on theory development, often using a combination of techniques such as interview, observation, and record review. On the other hand, the ethnographic approach may not be highly amenable to use of the Internet for data collection because it focuses on studying behaviors from within the culture.

Researcher expertise

In qualitative data collection the researcher plays a crucial role, and the successful use of the Internet for data collection will depend on how well the researcher can use interview skills within the context of Internet facilities available. In qualitative data collection the researcher is the instrument,[20] and so the expertise of the researcher is important. It takes a highly adept researcher to collect qualitative data of high quality because it is necessary to ask the right questions, clarify responses given, probe when responses need further exploration, and observe the respondent's behaviors within the context of the data collection session. The Internet may place certain restrictions on these necessary skills, particularly if computer facilities do not allow their full application.

The qualitative researcher who prepares for data collection over the Internet will not only need to know what information will be sought, but will need to understand the special opportunities and limitations afforded by Internet data collection. Therefore, the researcher is required to have adequate knowledge of the computer technology that is available for the study and will need to determine which approach to Internet data collection will be best, for example, real-time data collection with simultaneous written responses by the respondent ("instant messenger"), or using a set of interview questions that can be placed on an Internet site that will be responded to at a later time by subjects (discussion board). Chat groups and shared discussion boards are allowed by many computer applications, which make "virtual focus groups" possible.[8] No matter which strategy is chosen, the researcher's skill in developing questions for data collection that are salient to the study's focus and data needs are of utmost importance. Therefore, preparation of clear and understandable questions that are central to the topic is a crucial step in Internet data collection as it is in any qualitative data collection process. However, when the researcher develops questions for subjects to respond to at their convenience, this point is particularly important.

Real time versus later data collection

In today's highly technological society, it is possible to collect qualitative data over the Internet in real time, or synchronous communication, with the respondent providing responses verbally and in writing. On the other hand, the investigator may prefer to prepare open-ended questions to which the respondent would provide answers over the Internet in written form at a later convenient time or through asynchronous communication. Each of these approaches has advantages and limitations.

When data are collected over the Internet in real time, open-ended questions can be submitted to the Web site in advance or at the time that the computer interaction occurs. In either case the interviewer and respondent will need to schedule a time for the subject to respond in real time so that both the researcher and the respondent are available simultaneously. As the respondent types in the responses to the questions, the responses appear on the interviewer's screen instantly. The advantage of this approach is that the interviewer can clarify confusing written comments or probe comments that require further explication at the time the respondent supplies them. This can increase the quality of data and the interviewer has more control

over the data collected, such as keeping the respondent engaged and on the topic, thus enhancing the possibility that answers are provided to probes or clarifying questions.

Data can be collected in written form without voice capabilities. Verbal interaction between the interviewer and respondent is also possible when microphones and appropriate software are available. When real-time voice capabilities are used, more sophisticated computer equipment and programs are required. This means that the cost of data collection will be higher than without voice capabilities. In this approach, the researcher must have a live microphone. If there is a need for verbal interchange between the researcher and respondent, then the respondent also must have a microphone. In either case the respondent's computer system must have an active speaker so that probes and clarifying questions from the researcher can be heard. In addition, real-time computer data collection requires a fast Internet connection and fast computer, thus limiting the potential number of study participants.[8]

When Internet data collection is not conducted in real time, the researcher must review the respondent's comments after they have been entered at the Web site. The researcher will be required to check the Web site frequently so that clarifying questions and additional probes can be provided in response to the subject's answers to the original set of open-ended questions. Hence, probes and clarifying questions are submitted at a later time. In this circumstance, it is easy for the respondent not to respond at all or to provide partial answers to the information requested. It is also easier for respondents to wander off the topic rather than providing information relevant to the study. Therefore, the researcher has less control over the data collection process.

Internet data collection of qualitative data has an advantage over telephonic interviews because it costs less to connect to the Internet than making contact via telephone, particularly when subjects are from other states or geographic regions. Clearly, an advantage of Internet data collection is that it can afford an investigator access to respondents who would not have been available for recruitment locally since respondents can be recruited nationally or internationally, if appropriate, with little additional costs. However, when respondents are recruited over the Internet, demographic and screening data must be carefully obtained to ensure that respondents are part of the desired target population. In addition, the investigator needs to consider whether geographic location and culture could be an important factor in the nature of the experience communicated by the respondent. If so, the geographic region for recruitment would need to be restricted or the investigator should be careful that interpretations of the data reflect regional or cultural concerns and issues.

In the migraine study, qualitative data were collected through a series of discussion boards that were created as participants were recruited into the study and completed the on-line questionnaires. Each discussion board had a single login name and password, which was provided to all the participants for that group. Each board, which was planned to be the on-line equivalent of focus groups, consisted of between 4 and 8 participants. We felt that these numbers were adequate to encourage interaction and communication regarding experiences with migraine headaches by the participants while also being small enough to prevent the discussion boards from becoming overwhelming for participants who responded to the other participants' comments as well. The boards were not conducted in real time; participants logged in and out at their convenience over a period of between 3 and 5 weeks. The moderator/researcher initially posted a welcome message for each board, along with several questions to act as beginning "threads" to begin the discussion. Participants were then free to respond to the moderator's questions, begin new threads of their own, or merely read the other postings and post no response at all. Several advantages accrued as a result of this Internet approach. It was possible to have an ongoing discussion board/focus group with women from diverse geographical locations. Women

who otherwise could not have found a common time to meet were able to participate at different times of the day and night. At the completion of the discussion board, data were entered immediately into a word-processing package for analysis, rather than having to be transcribed and edited. There were also several major disadvantages of using this approach to gathering discussion group data in non-real time. Participants sometimes carried on their own individual conversations, not responding to questions or probes from the moderator. Some participants enjoyed reading the other postings, but did not themselves post any, or very few, comments. Many participants had very hectic schedules and tended to forget to post their comments to the discussion board, requiring regular reminders from the researchers.

DISCUSSION

Using the Internet in data collection has many exciting possibilities, and like any new method, will require creativity and planning to take advantage of the possibilities the technology offers, while at the same time minimizing threats to the collection of valuable data. However, there are some general guidelines that we believe will be helpful to the researcher who is beginning to incorporate the Internet into a methodological toolbox.

1. It is helpful to keep fairly tight control of the study, carefully screening potential participants to verify their appropriateness for the study, and providing individual passwords that can be used only to a limited extent in the study.
2. It may also be helpful to consider mixing use of the Internet with other approaches to communication. For example, screening might be done by phone to establish a relationship with the researcher, also again diminishing the possibility of fraud. In a study that spans several weeks, or months (as ours did), it may be helpful to occasionally "check in" with participants by phone to see how they are doing with the computer work.
3. Handling interviews or discussion groups on-line is rewarding and may be convenient for the researcher as well as the participant. However, we found that this experience was time-consuming for the researchers and required a major time commitment from the research staff and the participants.
4. It is important to have computer consultation readily available, unless the researchers themselves are expert in computer usage. Glitches do occur (as when a questionnaire will not advance to the next page, or error messages appear for no apparent reason), and someone needs to address them quickly to minimize participant frustration and study dropouts.
5. Use of public facilities, such as library computers, represents a feasible way for individuals to participate in Internet research even if they lack the usual access. As noted above, although this is feasible, it requires a greater commitment by researchers and participants.

Internet data collection opens up many opportunities for compiling data for research studies that can allow additional access to participants. However, Internet data collection has its limitations as well as advantages. The researcher needs to carefully consider the purposes of the study, the nature of the target population, skills of the investigative team, and available resources when deciding the use the World Wide Web for research data collection.

REFERENCES

1. Im E, Chee W. Issues in Internet research. *Nurs Outlook*. 2003;51:6–12.
2. Joinson A. Social desirability, anonymity, and Internet-based questionnaires. *Behav Res Methods Instrum Comput*. 1999;31:433–438.
3. Kelly MW, Oldham J. The Internet and randomized

controlled trials. *Int J Med Inf*. 1997;47:91–99.
4. Wagner TH, Patrick DL, Galer BS, Berzon, RA. A new instrument to assess the long-term quality of life effects from migraine: development and psychometric testing of the MSQOL. *Headache*. 1996;36:484–492.
5. McHorney CA, Ware JE, Raczek AE. The MOS 36—Item Short-Form Health Survey (SF-36); II: Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993;31:247–263.
6. Senior C, Smith M. The Internet . . . a possible research tool? *Psychologist*. 1999;12:442–444.
7. Sheehan KB, McMillan SJ. Response variation in e-mail surveys: an exploration. *J Advertis Res*. 1999:39(4):45–54.
8. Lakeman R. Using the Internet for data collection in nursing research. *Comput Nurs*. 1997;15:269–275.
9. Moloney MF, Melby MA. Trading tampons for pills: the lived experience of perimenopausal women with migraines. In: Proceedings of the society for Menstrual Cycle Research; June 7–9 2001; Avon, Conn.
10. Smith MA, Leigh B. Virtual subjects: using the Internet as an alternative source of subjects and research environment. *Behav Res Methods Instrum Comput*. 1997;29:496–505.
11. Houston JD, Fiore DC. Online medical surveys: using the Internet as a research tool. *M.D. Comput*. 1998;15(2):116–119.
12. Szabo A, Frenkl R. Considerations of research on Internet: guidelines and implications for human movement studies. *Clin Kinesiol*. 1996;50(3):58–65.
13. Stanton JM. An empirical assessment of data collection using the Internet. *Pers Psychol*. 1998;51:709–725.
14. Strickland OL. Challenges in measuring patient outcomes. *Nurs Clin North Am*. 1997;32:495–512.
15. Fleitas J. Computer monitor. Spinning tales from the World Wide Web: qualitative research in electronic environment. *Qual Health Res*. 1998;8:283–292.
16. Strickland OL. The practical side of measurement. *Behav Meas Lett Behav Meas Database Serv*. 1999;6(1):9–11.
17. Fawcett J, Buhle EL. Using the Internet for data collection: an innovative electronic strategy. *Comput Nurs*. 1995;13:273–279.
18. Wilmoth MC. Computer networks as a source of research subjects. *West J Nurs Res*. 1995;335–337.
19. Morse JM, Field PA. *Qualitative Research Methods for Health Professionals*. Thousand Oaks, CA: Sage; 1995.
20. Burns N, Grove SK. *The Practice of Nursing Research: Conduct, Critique and Utilization*. 4th ed. Philadelphia: Saunders; 2001.