Quality-of-Life Research on the Internet: Feasibility and Potential Biases in Patients with Ulcerative Colitis

ROY M. SOETIKNO, MD, RAMZI MRAD, VICTORIA PAO, LESLIE A. LENERT, MD

Abstract
Objective: The World Wide Web (WWW) is a new communications medium that permits investigators to contact patients in nonmedical settings and study the effects of disease on quality of life through self-administered questionnaires. However, little is known about the feasibility and, what is more important, the validity of this approach. An on-line survey for patients with ulcerative colitis (UC) and patients whose UC had been treated with surgical procedures was developed. To understand how patients on the WWW might differ from those in practice and the potential biases in conducting epidemiological research in volunteers recruited on the Internet, post-surgery patients who responded to the WWW survey were compared with those in a surgical practice.

Setting: The Internet and private practice surgical clinic.

Main outcomes: Scores from the Short form 36 (SF-36) Health Assessment Questionnaire and the Self-Administered Inflammatory Bowel Disease Questionnaire (IBDQ).

Results: Over a 5-month period, 53 post-surgery patients enrolled in the Internet study; 47 patients from a surgical clinic completed the same computer-based questionnaire. Surgically treated patients on the WWW were younger than their clinic counterparts (median age category 35–44 years vs. 45–54 years, p = 0.01) but more ill with a lower summary IBDQ score (168 vs. 186, p = 0.019) and lower health status across almost all dimensions of the SF-36 (p = 0.016).

Conclusions: It is feasible to conduct epidemiological research on the effects of UC on quality of life on the Web; however, systematic differences in disease activity between volunteer patients on the WWW and “in the clinic” may limit the applicability of results.

The explosive growth of Internet technology and of Internet users over the past 2 years has created a new virtual public meeting place. Many commercial entities have taken advantage of the millions of persons passing through this meeting place to provide services and sell products. Researchers have generally used this virtual meeting place to present and exchange research data among themselves. We believe, however, that it may also be possible to actually conduct research on the World Wide Web (WWW) and that this approach could both complement and, in some instances, replace the classic methods of conducting epidemiological research, particularly for studies of patients’ quality of life. By opening Web sites dedicated to clinical research and attracting patients to those sites, investigators might learn much about how different diseases affect patients’ quality of life in the community. However, before a meaningful study using patients on the WWW can be undertaken, important questions must be answered about the feasibility and validity of this approach. This paper describes
work on a cross-sectional study designed to evaluate the feasibility of conducting epidemiological studies in patients on the Internet.

The issue is not whether the data from health assessment questionnaires administered by computer over the WWW is reliable. The use of computers to directly collect health assessment data from patients is a well-established technology that has been shown to produce internally consistent responses when administered over the WWW. In some circumstances computer surveys have even been shown to have advantages over face-to-face interviews. Rather, the question is whether research results from questionnaires posted on the Web are generalizable. One set of problems is posed by the demographic differences between WWW users and the general U.S. population. While there are currently major differences, access to the WWW appears to be rapidly disseminating through the U.S. population. For example, the proportion of the U.S. and Canadian population who accessed the WWW in the previous 3 months increased by more than 50% between August of 1995 and March of 1996, from 8% to 13%. As WWW access devices get less expensive and are more closely integrated with television technologies, it is likely that WWW-based communications technologies will become as ubiquitous as television sets or even telephones.

A more important issue is whether WWW patients differ in some systematic way from patients in the community. WWW research will often have to rely upon self-selected patient volunteers. Volunteer patients can differ from general population of patients in unpredictable ways. "Surfing" the WWW is fundamentally an information gathering activity. Patients on the WWW may have some factor that motivates their search for information. One potential motivating factor might be a high level of disease symptoms. If disease activity is higher in WWW patients, this might limit the generalizability of certain types of epidemiological research conducted on the WWW. In this study, we investigate the feasibility of performing quality of life surveys in a specified patient population over the WWW and compare the health status of volunteer patients recruited over the WWW to volunteer patients recruited from a surgical clinic.

**Clinical Background**

Ulcerative colitis (UC) is a chronic inflammatory disease of the colon of unknown etiology. Its onset is in early adulthood, and the disease currently affects approximately 250,000 Americans. It is estimated that 18% of patients with UC have flare-ups every year; 57% have intermittent relapses; and only 25% are in remission. The life-long morbidity associated with UC is reflected in the approximately 250,000 visits and 36,000 hospital discharges per year for the management of UC and its complications. Furthermore, patients with long-standing and extensive UC have an increased risk of developing colon cancer, which may lead to further morbidity and early mortality.

Because of the chronic morbidity associated with the disease, quality of life is extremely important for UC patients. The disease often begins in early adulthood, and it is not uncommon for patients in their mid-thirties to have had more than 10 years of colitis and be confronted with a dilemma. They must decide whether to accept an increasing risk of colon cancer as well as periodic relapses of the disease or to undergo surgery to remove the diseased colon. The most widely used surgery for ulcerative colitis is a colectomy with ileal pouch-anal anastomosis (CIPAA). In this surgery, the colon is removed and the mucosa/submucosa of the rectum is replaced with the most distal aspect of the ileum that has been configured to form a pouch. This surgery allows patients to pass stool through their rectum, as opposed to requiring an ileostomy. While patients have frequent stools (five or six times a day under optimal conditions), they typically have control over their bowel functioning and often have minimal symptoms. A key issue in the use of this surgery is the quality of life for patients after the surgery. If it is near normal or even simply increases thanks to surgery, than patients with ulcerative colitis may find early surgery advantageous. If the quality of life of post-colectomy patients is far from normal health, UC patients may wish to delay surgery as long as possible and reduce cancer risks by other means (such as frequent colonoscopy to detect early cancers).

Surgery for ulcerative colitis is not frequently performed, and it is difficult to recruit large and representative populations of patients having had this procedure. We therefore attempted to locate and study patients who have had this procedure over the Internet. Because little is known about how representative patients on the Internet are of patients in general, we compared our patients recruited via the WWW with a modest-sized group of patients who had undergone a similar procedure recruited from a surgical clinic.

**Methods**

We constructed the WWW site using HTML 3.2, and we linked it to the Internet through WebSTAR 2.0 server software (StarNine Technologies, Marina Del Rey, CA). In order to automatically transfer patients’ data to a database, we used a program called WEBFM.
Figure 1 Design of the WWW site used for the study. Each box represents a series of WWW pages with a specific function. The sequence of travel through the WWW site is controlled by a finite state machine. Patients must complete a survey in the specified order but, as shown, can restart the survey at a later time through a special WWW page, and they can resume the survey where they left off.

A schematic diagram of the content of the WWW site is shown in Figure 1. An initial page described the survey and requested patient assistance (Fig. 2). This page had links to other pages which described rewards offered to WWW patients who participated in the study. The subsequent pages of the WWW site described the purpose of the study in detail, provided previews of the instruments with which we would be collecting data, and described other aspects of the research project. Links within the site allowed patients to view the results of a previous study of ulcerative colitis patients on the WWW and to learn more about the investigators behind the study. When patients finished reviewing this content and had learned as much as they wished about why we were undertaking the survey and what they would be asked, they followed a hyperlink to enroll in the study. At this point, the WWW site obtained informed consent from patients prior to their participation in the survey. Procedures to obtain informed consent were approved by the Stanford University Non-Medical Human Subjects Committee. To obtain informed consent, the patient entered his or her name at the bottom of an electronic version of our standard institutional consent form. The patient received back from the database a unique identifier (a username and password) with which he or she could log on to the survey. Logging on to the survey was assumed to mean that a patient had consented to participate in the study.

Questionnaire Elements

The survey first collected contact information that included an address, phone number, and—most important—an electronic mail address. It also collected data on the person enrolling in the study. Enrollees could specify that they were either patients or visitors. Enrollees who specified they were patients indicated the type of disease they had (ulcerative colitis or Crohn’s disease) or the type of surgical procedure (CIPAA or ileostomy) they had undergone. We then collected information on demographics: age, sex, marital status, number of children, education, occupation, and income. Completion of the “income” field was optional.

Next, patients completed an on-line version of the SF-36 Health Assessment Questionnaire, a standard health assessment instrument that measures quality of life in eight dimensions of health: physical health, mental health, emotional role-functioning, physical role-functioning, social health, general health perceptions, vitality, and pain. The maximum score that can be obtained on any of the dimensions of the SF-36 is 100 points.

Patients then completed the Self-Administered Inflammatory Bowel Disease Questionnaire (IBDQ). The IBDQ is a disease-specific health status instrument designed to measure quality of life in inflammatory bowel diseases such as ulcerative colitis and Crohn’s disease. The IBDQ evaluates four aspects of the effects of inflammatory bowel diseases on patients’ quality of life: physical health, systemic symptoms, emotional health, and social functioning. The IBDQ has been shown to be responsive to changes in intensity of inflammatory bowel diseases and is used in clinical trials. Previously published results de-
scribe typical levels of symptoms in normal patients, in typical clinic patient populations, and in Crohn's disease patients undergoing exacerbations of their disease. The maximum score that can be obtained on the IBDQ is 224 points for the combined scale, which is the simple sum of the scores on the four sub-scales. (Higher scores indicate a better quality of life.)

After completion of these two standard health assessment questionnaires, the survey measured patients' preferences for their current health using a visual analog scale (VAS) rating device and the Standard Reference Gamble. In the VAS rating task, patients used an animated scale to rate their quality of life on a scale anchored on perfect health at the top and death at the bottom. By clicking on the scale, the level moved up and down; simultaneously, the percent of the distance between perfect health and death was displayed numerically on the screen. Before submitting and recording a patient's answer, the program asked patients to confirm their ratings, using the question: "Your answers suggest that you view your current health as x% of perfect health. Is this what you mean? If so click on Go On." Patients then proceeded to the standard gamble rating task.

In the SG rating task, patients used an animated graphic to identify the maximum risk of death they would be willing to take to trade their current health for "perfect" health. The metric search for the maximum risk that a patient would take uses a Ping-Pong search grid that varied from 1% risk to 99% to 2% to 98% to 10 to 90% to 20% to 80% and so forth. After finding the lowest risk a patient would not take, and the highest risk he or she would take, to trade their current symptoms for perfect health, the program asked patients to confirm their response with the question, "Your answers suggest that you would be willing to accept a risk of death between x% and x + 5% to be cured of your current health problems. Is
this what you mean? If so, click on the text on the screen labeled Go On.” The procedure was virtually identical to that used in previously validated standalone computer elicitation instruments.15–17

The third part of the survey was a multimedia description of the effects of CIPAA surgery on quality of life. In this segment, previously recorded patients described in their own words how CIPAA surgery had affected various aspects of their quality of life. Topics were taken from content items of the IBDQ scale. Patients who had had CIPAA or other surgical procedures to treat UC rated the similarity of their health condition to that described by the patients. The results of this comparison are beyond the scope of this paper.

Patient Recruitment

To recruit patients on the Internet, we requested that the study Web site be added to the WWW-page databases of five commercial Internet search engines (Yahoo, Lycos, AltaVista, WebCrawler, and Infoseek). We also requested hypertext links to our WWW site at the WWW site of the Crohn’s-Colitis Foundation of America. In addition, we posted an announcement of our study to Internet newsgroups devoted to patients with ulcerative colitis and postoperative patients. Patients recruited over the WWW were offered a choice of a small gift (a Stanford University mug, key chain, or pen) as compensation.

We recruited clinic patients who had undergone CIPAA procedures by identifying all patients who had received the procedure between 1971 and 1995 from the clinic’s records. Of these patients, we contacted patients with telephone numbers in the San Francisco Bay area and asked them to come in to our laboratory to take the same questionnaire. CIPAA patients were offered $10 compensation. Clinic patients completed the survey in the investigator’s offices. They were logged into the study’s WWW site by a research assistant. Patients who were unfamiliar with use of the computer were taught how to use it to respond to the questions in the survey. Patients then went through the WWW site on their own; however, they were monitored by the research assistant and were offered whatever technical assistance they needed to complete the survey.

Missing Data

We periodically identified the Internet patients with incomplete or partial responses to the survey by searching the database. We sent each patient two electronic mail reminders at approximately 1-month intervals asking patients to complete the survey. The reminder notices explained the purpose of the survey and the importance of obtaining complete data on all patients. The letter also reminded patients of their usernames and passwords, and it gave them complete instructions on how they could log in again to finish the parts of the survey that they had left incomplete.

Statistics

All responses were automatically recorded into an electronic database. Calculations of quality of life ratings were performed within the database. Results were exported from the database, and statistical tests were conducted using JMP 3.1 (SAS Institute, Cary, NC) for Macintosh computers.

Demographic features of the two groups were compared using Chi-square tests. Differences in quality of life as measured by the IBDQ and the SF-36 were evaluated using a simultaneous Multiple Analysis of Variance (MANOVA) test of the subscales of each health index, with Internet use as the independent variable. Quality of life ratings for individual sub-scales were compared using paired t-tests. Quality of life ratings for the visual analog scale ratings and Standard Gamble utility values were compared using the Wilcoxin sign-rank test. The internal validity of health status measurements was evaluated by assessing the partial correlation among sub-scales of the SF-36 or IBDQ. The validity of utility elicitation was assessed by using multiple linear regression to model the relationship between the SF-36, the IBDQ, and preference ratings for patients’ current health.

Results

The WWW site was opened on August 2, 1996, and requests for indexing and listing of links were sent during the first month of operation. Between August 2 and January 4, 1997, the patient recruitment Web page was accessed 5,391 times. During this time, more than 204 ulcerative colitis patients, 51 CIPAA patients, 13 Crohn’s disease patients, 19 ileostomy patients, and 26 persons who identified themselves not as patients but as “visitors” enrolled in our study. About 5.8% of accesses of the study’s home-page led to eventual enrollment of a participant in the study. Participants were recruited from 14 different countries and 37 different states of the United States of America. Most foreign participants were from Canada, the United Kingdom, and Australia.

The remainder of this paper describes results with interviews of post-CIPAA patients recruited over the Web and the comparison group recruited from the surgical clinic.
For the comparison group, we recruited 47 post-CIPAA patients using the records of a single community-based surgical practice. Records showed that 149 patients had received the CIPAA procedure between 1971 and 1995. Of these 149, we were unable to contact 53; 30 patients declined to participate, and 19 were interested in participating but unable to come to our offices to be interviewed.

Of the 53 CIPAA patients recruited from the WWW, 52 completed the SF-36 and IBDQ portions of the survey. Completion rates for preference assessment ratings were also high: 79% of WWW patients completed the visual analog scale and standard gamble rating tasks. All 47 of the clinic patients completed both health status instruments, and 46 of 47 completed preference rating tasks. Of the 53 WWW patients, 46 provided electronic mail addresses (89%) so that we could send reminder notices to those who failed to complete the survey. Reminder notices had only a small effect on completion rates; two reminder notices increased the rate of completion of VAS and Standard Gamble rating tasks by about 8% over the base rate.

**Demographic Data**

Web patients were younger than clinic patients (p < 0.01, Pearson Chi-square), with a median age of between 35–44 years compared with 45–54 years for clinic patients. There was a similar proportion of female patients in both groups (47% for WWW patients and 42% for clinic patients, p = 0.42, Pearson Chi-square). Approximately 50% of WWW users and 70% of clinic patients were married (p = 0.083). Clinic patients were more likely to report having children (75% vs. 46%, p < 0.01, Pearson Chi-square). Both groups of patients were highly educated: 97% of WWW patients and 88% of clinic patients had some college education (p = 0.07). Clinic patients reported slightly higher incomes than WWW patients (median $60K to $80K per year vs. $40K to $60K per year (p = 0.12, Pearson Chi-square).

**IBDQ Scores**

There was no statistical evidence of an association between patients’ summary IBDQ scores and their age, sex, or income levels. Therefore, we did not adjust for differences in demographics between the two groups in comparisons. The summary IBDQ scores were near normal (mean 186) in clinic patients and moderately depressed in WWW patients (mean 168, p = 0.0195 vs. clinic patients.) Distributions of summary scores are shown in Figure 3. WWW patients have a wider range of symptom scores in addition to lower mean scores. Results show that a large portion of WWW patients (about 25%) have severely symptomatic disease (less than 145 (12)); whereas only a few clinic patients had such severe symptoms. Analysis of the IBDQ by subscale confirmed that WWW users had lower health in general (p = 0.02, Hotelling–Lawley Trace—an approximation of the F statistic for multiple observations within an individual). However, the degree of impairment differed among sub-scales (p = 0.038). The difference seen in IBDQ summary scores was the result of differences in symptoms for only two of the four scales (the Bowel symptom scale and the Emotional Health scale). Values for the Systemic Symptoms scale and the Social Health scale were nearly equal among members of the two groups.

**SF-36 Health Questionnaire**

Similar effects were seen in the more general SF-36 health status measure. As shown in Figure 4, the health status of WWW patients was lower in almost all dimensions of health than in clinic patients (p = 0.016, Hotelling–Lawley Trace). There were important differences in health in all dimensions of the SF-36 except for the physical health scale. While the
IBDQ scale primarily showed differences in bowel symptoms and emotional health, the largest differences in the SF-36 subscales were seen in social health and Role-emotional health scales. Overall, SF-36 scores suggested that the health of WWW patients was similar to that of general medical patients in the Medical Outcomes Study and that the health of clinic patients was somewhat better.

**Internal Validity of Health Assessment Questionnaires**

To evaluate the internal validity of the IBDQ and SF-36 questionnaires administered in HTML format, we examined the partial correlation among sub-scales of each index. IBDQ sub-scale scores had good evidence of internal validity, with a partial correlation among dimensions ranging from 0.60 to 0.81. SF-36 sub-scale scores also showed evidence of internal validity. Partial correlations between dimensions ranged from 0.29 to 0.69, with the lowest correlation being between the Mental Health scale and both the Role-emotional scale and the Physical Health scale. The highest correlations were observed between the Physical Health scale and the Pain, Social Health, and Vitality scales.

**Utility Elicitations**

Approximately 80% of WWW patients completed the preference elicitation tasks. Many users indicated in electronic mail communications to us that they could not complete this task because of technical difficulties. These difficulties included improperly sized screens that hid navigation commands and browsers that were not compatible with the Javascript (Netscape Communications, Mountainview, CA) HTML extensions used in the survey. Visual analog scale ratings also suggested poorer health in WWW patients versus clinic patients (0.67 vs. 0.79, p = 0.03 Wilcoxon sign-rank test). However, there was no difference in Standard Gamble ratings for current health (0.829 WWW vs. 0.826 clinic, p = 0.98) among the two groups. We tested for potential biases induced by drop-outs from the study by comparing IBDQ summary scores in WWW patients who did and did not complete the survey. Patients who did not complete the survey had IBDQ scores similar to those who did (167 vs. 172, respectively, p = 0.55).

**Internal Validity of Utility Elicitations**

VAS and Standard gamble ratings for current health were uncorrelated (r = 0.006). We used linear regression to evaluate the relationship between both VAS and Standard Gamble utility values and IBDQ and SF-36 health status measures. There was a fairly high correlation between the summary IBDQ score and utility ratings (r = 0.51 for the VAS ratings and r = 0.41 for the SG). Linear models using patient scores on various dimensions of SF-36 to predict preference ratings explained a large proportion of the variance in VAS preference ratings (r = 0.58), but only a moderate amount of the variance in Standard Gamble ratings (r = 0.34). Combining SF-36 health indices and the IBDQ summary score in regression analysis did not materially improve the model's ability to explain variation in VAS or Gamble preference ratings.

**Discussion**

The WWW is a virtual public meeting and marketplace where researchers can easily open a "storefront" and recruit patients to participate in clinical studies. Our results show that it is possible to enroll large numbers of patients (more than 250) in a short period of time (5 months) in an Internet study. Conducting research on the WWW requires some computer infrastructure. While it is possible to post a simple form on the WWW without software to monitor data quality, we feel it would be unlikely that a survey would generate valid data. We designed our WWW survey with the goal of convincing patients that we were conducting a serious research effort. We paid great attention to the graphics and page layout of the survey, and we used recorded voice instructions, a relatively rare technology on the Internet when the study started. Both these methods increased the clarity of the Web site and served as qualitative arguments to convince respondents that we truly cared about the responses they would provide.

Before any patient could participate in the study, he or she had to view a description of the purpose of the study and provide consent for participation using standard forms. Most WWW patients (more than 90%) were willing to provide either a phone number or electronic mail address for follow-up purposes. One potential problem with automated patient data entry might arise due to visitors to the site impersonating patients. While we cannot exclude this possibility, we specifically allowed visitors to identify themselves as such so that interested parties could inspect our site without contaminating the experiment. Furthermore, we believe that there was very little incentive for a non-patient to fill out a complete survey with false information. One way to confirm the diagnoses of patients would be to request their medical records. In a previous survey, focused on evaluating the willingness of UC patients to participate in clinical research over the Internet, we found that more than 70% of UC patients were willing to release medical records to allow researchers to confirm their diagno-
Future research should examine the accuracy of self-reported diagnostic information in WWW surveys.

We had to decide what to do about “missing” responses. In developing an interface design for questionnaires, we kept in mind that patients might tire of reporting responses in a long questionnaire and want to move on to the next part of the study to “see what was there.” This behavior is consistent with the typical use of the WWW (namely, “browsing” for information), and we anticipated that this might threaten the validity of the results of our study. Most commercial technologies for processing forms sent to a database over the Internet are poor in handling missing data; they simply reject forms with missing data and often return a new blank form that the user has to complete. This is often frustrating for WWW users. To avoid this problem, we developed software that dynamically built custom-tailored WWW pages with just the questions that patients had forgotten to answer. As a result, we had essentially no missing data for health assessment questionnaire items. Missing data occurred only when patients terminated participation in the study, and all data items for an individual were complete up to the point at which he or she declined to further participate. We did not modify existing standard questionnaires to allow patients to refuse to respond to single items because of our hesitancy to tamper with “standardized” health assessment questionnaires such as the SF-36 and the IBDQ. Future work will examine the importance of this option and its effects.

The completion rates for the SF-36 and IBDQ were very high. The survey instrument in this study had more than 80 multiple choice questions. The results show that it is possible to undertake “serious” research efforts in patients on the WWW. However, there also appeared to be limits to the endurance of WWW patients; parts of the survey that followed our two health assessment questionnaires and which used advanced Internet technologies such as frames, sound, and HTML scripting languages had lower completion rates. This is likely to be due in part to fatigue, to the technologies outpacing the computer equipment of the patients, and to the greater difficulty of value measurement questions in these sections. “Mass” mailing electronic reminder notices to complete surveys had minimal effects on completion rates. Further work is needed to evaluate the effectiveness of a more personalized approach.

The demographics of WWW patients who responded to this survey were similar to those we observed in a previous study of ulcerative colitis patients on the WWW.18 In this study and in our previous study, we observed equal proportions of male and female patients (a higher proportion of females than is typically found in surveys of WWW users). While WWW patients were well educated, and levels of education were typical for that reported by WWW users, income was lower than that reported for Web users. Patients on the WWW were clearly demographically distinct from clinic patients; they were younger (median age 35–44 years) and more frequently single and without children.

In this study, demographic factors were not associated with symptoms as measured by the IBDQ scale. Therefore, it is unlikely that differences in health status measurements are due to differences in the demographic make-up of the two groups. Rather, it seems likely that volunteer patients from the WWW who responded to our survey actually had more symptoms than volunteer patients from the surgical clinic.

Data from the IBDQ and the SF-36 suggest different sorts of problems. Differences in IBDQ subscales were primarily in the level of bowel symptoms (for example, the number of bowel movements, leakage, flatulence, etc.) and in the emotional health of CIPAA patients (for example, bowel problems leading to sexual dysfunction and other emotional problems). Systemic symptoms and social functioning were roughly equal. The SF-36 health data confirmed that WWW patients had problems with emotional health, role functioning (both physical and emotional), pain, vitality, and general perceptions. However, physical health scores were similarly high.

While we did observe a significant difference in VAS ratings for current health among WWW users and clinic patients, we did not in the Standard Gamble utility elicitation. The failure to detect differences in preferences using the Standard Gamble may be due to this method being less sensitive to changes in quality of life or due to the differences not being clinically important to patients. Alternatively, both the SF-36 and the IBDQ use a time frame that focuses on the patient’s recent experience (past 4 weeks and past 2 weeks, respectively), whereas the Gamble rating task asks patients to integrate their disease experiences over their lifetime. It could be that patients’ overall life experiences are similar and that only recent experiences with the activity of disease are different. If patients are aware of the waxing and waning nature of their disease, then they expect remission with time and medical treatment of their current acute symptoms and thus discount current disease activity in preference ratings.

Clinically, the health of many post-CIPAA patients on
the WWW appeared to be poor. While the median values for IBDOQ scores are similar to those of UC patients without active disease, fully one-quarter of patients had scores as severe as those seen at the nadir of a flair-up of Crohn’s disease. One explanation for the relatively poor health of CIPAA patients on the WWW is that they are driven by higher levels of symptoms activity to seek information and support relevant to their condition on the WWW. In this activity, they might have come across our WWW site and decided to respond to the survey. The link between disease activity and information seeking behavior in patients needs to be explored in greater detail in future research.

The difference in the health of the two populations of volunteers is likely to be due, in part at least, to differences in incentives for participation. WWW patients responded from home or work without interruption of their activities. Clinic patient had to physically come to a relatively distant university medical center to participate in the study (as is typical for many research projects). Some of the clinic patients did indicate to us that they were not feeling well enough to make trip to participate in the survey. However, other clinic patients also told us that they were simply too busy with their daily activities to participate. We feel that by moving the site of clinical research to the home and work place, through the use of the Internet, we are likely to get more generalizable results than using volunteers from medical settings despite the current differences in demographics between the Web and the Clinic.

The worse health of UC patients from the WWW observed in our survey stands in contrast to the findings described by Bell and Kahn. In an anonymous WWW survey that was designed to allow participants to benchmark their health using the SF-36, they found survey respondents to be healthier, on average, than patients in the Medical Outcomes study. The differences in the findings of these two studies is likely to be because our study focused on recruitment of a specific group of self-identified patients. The Bell and Kahn study targeted a general Internet audience interested in health assessment. It is likely that procedures for recruiting patients will have important effects on the specific findings from WWW-based studies. Linking SF-36 data collection to comparative feedback on health status (as was done with the Bell and Kahn study) is also a major change in how the SF-36 is administered that may change norms for responses. Further research is needed on the effects of this powerful incentive (or apparently powerful incentive, given the large number of patients recruited by Bell and Kahn) to participate in WWW studies.

If volunteer patients recruited on the WWW can be more ill or less ill than clinic patients, depending on the method they were recruited by, this may have important implications for the types of medical studies that can be carried out on the WWW. The WWW may be a better tool for performing longitudinal studies of quality of life than cross-sectional studies would be. Creating a panel of patients chosen to represent known factors influencing disease outcomes might be a valid way to study long-term effects of disease on quality of life. The ability to follow-up UC patients by electronic mail and to collect data from patients over the WWW might make this a very cost-effective method for studying patients’ outcomes.

A second option is to not attempt descriptive studies at all but to use the WWW for randomized trials. If UC patients on the WWW are in the midst of a flair of disease symptoms, and if this in fact underlies their search for information on the WWW, they may be an ideal population for conducting randomized trials of nonpharmaceutical strategies for disease management. Given the relatively large numbers of patients on the WWW, it may be a very useful medium to conduct studies of the decision support techniques, educational methods, or even behavior modification interventions that can be administered over the WWW. In a randomized trial, potential biases in populations would be controlled for and would be less important, although the demographics of the patients on the WWW might still limit the interpretation of results.

Conclusions

It appears to be feasible to conduct quality of life research in patients on the WWW; UC patients on the WWW appear to be both interested in participating in research and cooperative. Post-CIPAA patients on the WWW had higher levels of bowel-related symptoms and lower health status than a group of clinic patients. While there were demographic differences between the two groups, these differences did not explain measured differences in quality of life. Thus, while it is feasible to study patients’ quality of life on the WWW, the assessment of the overall validity of this approach and definition of its specific strengths and pitfalls will require further study. However, given low cost and favorable logistics, there is potentially much to be learned about the effects of disease on patients in the community from research in patients on the WWW.

We thank Ms. Lisa Matrich for her help in assembling the software used in this study, Dr. Donald Peck for providing access to his patients, and Manijeh Parineh for her help in adminis-
tering the survey. We also thank Dr. Dawn Provenzale for her helpful discussions of background material. Finally, we thank the patients who participated in this study for their patience and enthusiasm.

References

16. Lenert LA, Morss S, Goldstein MK, Garber AM. Measurement of the validity of utility elicitations performed by computer interview. Medical Care, in press.