



Ethical issues in qualitative research on internet communities

Gunther Eysenbach, James E Till

The internet is the most comprehensive electronic archive of written material representing our world and peoples' opinions, concerns, and desires. Physicians who surf the internet for the first time are often stunned by what they learn on websites set up by lay people or patient self support communities. Material on these venues can be a rich source for researchers interested in understanding the experiences and views of people and patients. Qualitative analysis of material published and communicated on the internet can serve to systematise and codify needs, values, concerns, and preferences of consumers and professionals relevant to health and health care. While the internet makes people's interactions uniquely accessible for researchers and erases boundaries of time and distance, such research raises new issues in research ethics, particularly concerning informed consent and privacy of research subjects, as the borders between public and private spaces are sometimes blurred.

Internet communities

Internet communities provide a way for a group of peers to communicate with each other. They include discussion boards on websites, mailing lists, chat rooms, or newsgroups. Examples of health related mailing lists can easily be found by inserting a key word such as "cancer" in the search box at the Topica website (www.topica.com/). One example is the Breast Cancer Mailing List, based at Memorial University in Newfoundland (www.bcmlist.org), which provides a forum for those seeking peer support and information, with an emphasis on information on treatment and disease, practical information (such as relevant online resources), personal experiences, and emotional support.

Qualitative research on the internet

Qualitative research seeks "to acknowledge the existence of and study the interplay of multiple views and voices—including, importantly, lay voices."¹ Internet postings are accessible for qualitative research of these voices—for example, to determine information needs and preferences of consumers or to investigate how health related information can best be converted into knowledge and disseminated widely.²

Three different types of internet based research methods can be distinguished.³ One is passive analysis, such as studies of information patterns on websites or

Summary points

Internet communities (such as mailing lists, chat rooms, newsgroups, or discussion boards on websites) are rich sources of qualitative data for health researchers

Qualitative analysis of internet postings may help to systematise and codify needs, values, and preferences of consumers and professionals relevant to health and health care

Internet based research raises several ethical questions, especially pertaining to privacy and informed consent

Researchers and institutional review boards must primarily consider whether research is intrusive and has potential for harm, whether the venue is perceived as "private" or "public" space, how confidentiality can be protected, and whether and how informed consent should be obtained

interactions on discussion groups without the researchers actually involving themselves. Examples include the study of helping mechanisms and content of online self help groups for colorectal cancer,⁴ breast cancer,⁵ Alzheimer's disease,⁶ and eating disorders.⁷ The second type of online research is through active analysis, in which researchers participate in communications—for example, to determine the accuracy of responses to healthcare questions on the Usenet.⁸ In the third type researchers identify themselves as such and gather information in the form of online semistructured interviews, online focus groups, or internet based surveys or use the internet to recruit participants for "traditional" research.

Internet communities' members do not expect to be research subjects

There is increasing evidence that researchers posting or "lurking" on such communities may be perceived as intruders and may damage the communities. King quotes a group member who, realising that the community had been monitored by a researcher,

Unit for
Cybermedicine and
eHealth,
Department of
Clinical Social
Medicine,
University of
Heidelberg,
Bergheimer Str 58,
69115 Heidelberg,
Germany

Gunther Eysenbach
senior researcher

Joint Center for
Bioethics, University
of Toronto, Toronto,
Ontario M5G 2M9,
Canada

James E Till
professor emeritus

Correspondence to:
G Eysenbach
ey@yi.com

BMJ 2001;323:1103-5

retreated from a mailing list with the remark: "When I joined this, I thought it would be a support group, not a fishbowl for a bunch of guinea pigs. I certainly don't feel at this point that it is a safe environment, as a support group is supposed to be, and I will not open myself up to be dissected by students or scientists."⁹

One subscriber to the Breast Cancer Mailing List responded to one of the frequent requests from researchers interested in obtaining insights into patients' personal experiences with breast cancer: "Why can't researchers do it the 'hard way' as they used to . . . and leave us alone on the Breast-Cancer list?"¹⁰ Sharf reported that among 14 people from a mailing list contacted to obtain informed consent, one woman was "somewhat hostile, assuming that [the researcher] had behaved voyeuristically, taking advantages of people in distress" and that "the idea of using the conversations as data had not occurred to many members."¹¹

We searched the Dejanews archive (<http://groups.google.com/>) to find comments of newsgroup participants responding to researchers' requests. Entering the search terms "research survey project health" identified 85 messages from researchers. By following the message threads, we could find newsgroup participants' reactions, which were often negative. Among the concerns expressed was that often "the researcher isn't familiar with newsgroup culture—problems akin to those occasionally experienced by anthropologists when they try to interpret the behaviour of cultures they really don't understand. And on the rare occasions when I've seen someone who's part of that culture attempt to study the group he/she is part of, it resulted in a rippling sense of resentment and betrayal among those who find such things underhanded."

Public spaces or private rooms?

Informed consent, privacy, and confidentiality are basic ethical tenets of scientific research on people.¹²⁻¹⁴ To determine whether informed consent is required, you first have to decide whether postings on an internet community are "private" or "public" communications. This distinction is important because informed consent is required "when behaviour of research participants occurs in a private context where an individual can reasonably expect that no observation or reporting is taking place."¹⁵ On the other hand, researchers "may conduct research in public places or use publicly available information about individuals (such as naturalistic observations in public places and analysis of public records or archival research) without obtaining consent,"¹⁵ and "research involving observation of participants in, for example, political rallies, demonstrations or public meetings should not require Research Ethics Board review since it can be expected that the participants are seeking public visibility."¹⁶

Although publication on the internet may have parallels to publishing a letter in a newspaper or saying something in a public meeting, there are important psychological differences, and people participating in an online discussion group cannot always be assumed to be "seeking public visibility." On the internet the dichotomy of private and public sometimes may not be appropriate, and communities may lie in between.

Several measures can be used to estimate the perceived level of privacy. Firstly, if a subscription or

some form of registration is required to gain access to a discussion group then most of the subscribers are likely to regard the group as a "private place" in cyberspace.¹⁷

Secondly, the number of (real or assumed) users of a community determines how "public" the space is perceived to be: a posting to a mailing list with 10 subscribers is different from a posting to a mailing list with 100 or 1000 subscribers. However, as messages sent to mailing lists are sometimes also stored in web accessible archives, the actual number of people accessing messages may be greater than assumed and may be impossible to determine.

Thirdly, and perhaps most importantly, the perception of privacy depends on an individual group's norms and codes, target audience, and aim, often laid down in the "frequently asked questions" or information files of an internet community. For example, SickKids is a discussion list for children who are ill. The information file about the mailing list states that "adults will NOT be permitted to participate on this list as its purpose is to provide kids with their own personal place to share." It seems clear that children who send messages to this list are unlikely to be "seeking public visibility." Similarly, a virtual self support group of sexual abuse survivors was reported to have a group policy explicitly discouraging interested professionals who were not sexual abuse survivors from joining the group,⁹ yet a decision was made to analyse postings without obtaining prior or retrospective consent from the group members.¹⁸

Can informed consent be waived?

If it is thought that a community may be perceived to be private, the next question is whether informed consent for passively analysing the postings is needed or whether this requirement can be waived. In clinical studies non-intrusive research such as retrospective use of existing medical records may be conducted ethically without the express consent of the individual subjects if the material is anonymised at the earliest possible stage, if there is no inconvenience or hazard to the subjects, and if the institutional review board has reviewed and agreed the research protocol.¹⁹ Similar considerations may be applied to passive analysis of messages on mailing lists. When considering potential hazards to group participants or the community as a whole, privacy issues are especially important, and it should be considered whether publication of the results (especially when mentioning the group name) may negatively affect group members or harm the community as a whole. Much will depend on what data will be collected and how they will be reported, how vulnerable the community or sensitive the topic is, and the degree to which the researcher interacts with group members.

How can informed consent be obtained?

If a researcher and the institutional review board feel that obtaining informed consent is necessary, how should it be obtained? For mailing lists, there are basically two possibilities. The first is to send an email to the mailing list describing the research prospectively and giving participants the opportunity to withdraw from the list. The second is to ask retrospectively each person whose postings have been or will be used, giving them the possibility to withdraw themselves from the analysis.

The first approach is intrusive and can be done only for those mailing lists or chat rooms where the entry of new members can be monitored, so that they can be briefed on the ongoing research. There is also a considerable danger that announcing the research may influence future communication patterns or provoke many members to opt out (which may damage the community). The second approach is much less intrusive, can ensure that researchers correctly interpret statements of participants, and may even avoid false stories of personal illness being analysed.²⁰ However, this approach is cumbersome and time consuming,¹¹ especially as email addresses may be no longer valid and many participants may not respond.

Obtaining permission from the "list owner" (the individual responsible for maintaining the mailing list) or moderator (if any) is rarely an adequate way for a researcher to obtain "community consent," as neither can properly claim to speak for all of the participants in a mailing list. However, they may know the online community better than the researcher, so they can be a useful starting point to find out what the group norms are and what would be the best way of obtaining informed consent from group members.

Privacy and confidentiality

The internet holds various pitfalls for researchers, who can easily and unintentionally violate the privacy of individuals. For example, by quoting the exact words of a newsgroup participant, a researcher may breach the participant's confidentiality even if the researcher removes any personal information. This is because powerful search engines such as Google can index newsgroups (groups.google.com), so that the original message, including the email address of the sender, could be retrieved by anybody using the direct quote as a query. Participants should therefore always be approached to give their explicit consent to be quoted verbatim and should be made aware that their email address might be identifiable. Another reason why researchers should contact individuals before quoting them is that the author of the posting may not be seeking privacy but publicity, so that extensive quotes without attribution may be considered a misuse of another person's intellectual property.

Conclusions

The main problem with using internet communities for research is that researchers may have difficulty separating spaces regarded as private from spaces regarded as public. We recommend that researchers and institutional review boards should carefully consider these points when developing and reviewing research protocols and should involve members of the group they want to study in these considerations (see box). Our examples show how complex such considerations may be, being further compounded by our lack of knowledge about the psychological aspects of cyberspace, the risks and benefits of internet based research, and the technical complexities involved. Best practice guidelines for internet research and compilations of case studies, from the perspective of both researchers and internet communities, are needed.

Proposed considerations for researchers and institutional review boards before studying an internet community*

Intrusiveness—Discuss to what degree the research conducted is intrusive ("passive" analysis of internet postings versus active involvement in the community by participating in communications)

Perceived privacy—Discuss (preferably in consultation with members of the community) the level of perceived privacy of the community (Is it a closed group requiring registration? What is the membership size? What are the group norms?)

Vulnerability—Discuss how vulnerable the community is: for example, a mailing list for victims of sexual abuse or AIDS patients will be a highly vulnerable community

Potential harm—As a result of the above considerations, discuss whether the intrusion of the researcher or publication of results has the potential to harm individuals or the community as a whole

Informed consent—Discuss whether informed consent is required or can be waived (If it is required how will it be obtained?)

Confidentiality—How can the anonymity of participants be protected (if verbatim quotes are given originators can be identified easily using search engines, thus informed consent is always required)

Intellectual property rights—In some cases, participants may not seek anonymity, but publicity, so that use of postings without attribution may not be appropriate

*Researchers should explicitly address these issues in their research protocol

All authors of the quotes from newsgroups and mailing lists have given their consent to quote them.

Contributors: Both authors contributed equally to this manuscript.

Funding: JET's contributions were partially supported by an award from the National Cancer Institute of Canada.

Competing interests: None declared.

- 1 Barbour RS. The role of qualitative research in broadening the 'evidence base' for clinical practice. *J Eval Clin Pract* 2000;6:155-63.
- 2 Eysenbach G. Consumer health informatics. *BMJ* 2000;320:1713-6.
- 3 Eysenbach G, Wyatt JC. Facilitating research via the internet. In: McKenzie B, ed. *Internet and medicine*. Oxford: Oxford University Press (in press).
- 4 Klemm P, Reppert K, Visich L. A nontraditional cancer support group. The internet. *Comput Nurs* 1998;16:31-6.
- 5 Sharf BF. Communicating breast cancer on-line: support and empowerment on the internet. *Women Health* 1997;26(1):65-84.
- 6 White MH, Dorman SM. Online support for caregivers. Analysis of an internet Alzheimer mailgroup. *Comput Nurs* 2000;18:168-76.
- 7 Winzelberg AJ. The analysis of an electronic support group for individuals with eating disorders. *Comput Human Behav* 1997;13:393-407.
- 8 Seaboldt JA, Kuiper R. Comparison of information obtained from a Usenet newsgroup and from drug information centers. *Am J Health Syst Pharm* 1997;54:1732-5.
- 9 King SA. Researching internet communities: proposed ethical guidelines for the reporting of results. *The Information Society* 1996;12(2):119-28.
- 10 Till JE. Research ethics: internet-based research. Part 1: On-line survey research. <http://members.tripod.com/~ca916/index-3.html>. (updated 18 Nov 1997, accessed 20 Jan 2001).
- 11 Sharf BF. Beyond netiquette. The ethics of doing naturalistic discourse research on the internet. In: Jones S, ed. *Doing internet research*. London: Sage; 1999:243-56.
- 12 World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. www.wma.net/e/policy/17-c_e.html (update 7 Oct 2000, accessed 20 Jan 2001).
- 13 Jones RA. The ethics of research in cyberspace. *Internet Res* 1994;4(3):30-5.
- 14 Frankel MS, Siang S. *Ethical and legal issues of human subjects research on the internet—report of an AAAS workshop*. Washington, DC: American Association for the Advancement of Science, 1999.
- 15 American Sociological Association. American Sociological Association code of ethics. www.asanet.org/members/ecoderev.html (updated 1 Aug 1999, accessed 12 Jan 2001).
- 16 Social Sciences and Humanities Research Council of Canada. Tri-Council policy statement: ethical conduct for research involving humans. www.sshrc.ca/english/programinfo/policies/index.htm (updated 14 Sep 2001).
- 17 Mayer M, Till JE. The internet: a modern Pandora's box? *Qual Life Res* 1996;5:568-71.
- 18 Finn J. An exploration of helping processes in an online self-help group focusing on issues of disability. *Health Soc Work* 1999;24:220-31.
- 19 Royal College of Physicians Committee on Ethical Issues in Medicine. Research based on archived information and samples. Recommendations from the Royal College of Physicians Committee on Ethical Issues in Medicine. *J R Coll Physicians Lond* 1999;33:264-6.
- 20 Feldman MD. Munchausen by internet: detecting factitious illness and crisis on the internet. *South Med J* 2000;93:669-72. (Accepted 4 October 2001)