

Assessing an Internet health information site by using log analysis: the experience of the National Cancer Institute of Brazil

*Paulo R. Vasconcellos-Silva,¹
Luis D. Castiel,² and
Francisco J. U. Rivera²*

Key words: Internet, public health informatics, health education, mass media, neoplasms.

¹ National Cancer Institute of Brazil (INCA), Rio de Janeiro, Brazil. Send correspondence to: Paulo R. Vasconcellos-Silva, INCA, Praça Cruz Vermelha 23, 4° andar — Direção Geral, CEP 20230-130, Rio de Janeiro, Brazil; e-mail: paulor@inca.gov.br

² Oswaldo Cruz Foundation, National School of Public Health, Rio de Janeiro, Brazil.

In recent years there has been an information revolution of epic proportions. Many people have thought that new information technology would solve numerous problems with democracy and education because nothing since Gutenberg's invention of the printing press has provided the world with such potential for making huge amounts of information easily available at low cost to massive numbers of people. For example, information technology has become an integral part of the modern concept of public health and national health care policies (1).

Public health officials have advocated disseminating information through the Internet, television, and other mass media as a way to bring about health behavior changes, including discouraging such harmful habits as smoking, excessive sunbathing, doing too little exercise, and eating a diet high in fats. Nevertheless, while credible and relevant information is necessary to modify behaviors and produce lifestyle changes, it alone is not sufficient (2).

So why can't good information by itself bring about behavior changes? And why hasn't trustworthy public health information on the Internet had the same impact that commercial messages on the Internet have had? There are a variety of reasons. One of the most important is the tremendous volume of material that is available on the Internet. The number of health-related Web sites is now estimated to be more than 100 000. Making matters worse is the fact that only half of those sites have had their content reviewed by a doctor (3). With a lack of clear standards for quality and ethics for Internet health sites, there can be tremendous conflicts and contradictions between the information that various sources present (4).

The flood of information from the Internet and other communications media has sometimes diverted public interest away from the important and preventable causes of disease. Indiscriminate media attention on the more sensational aspects of diseases can lead the public to focus on minimal or unproven risks (5). For example, with cancer-related issues some persons have responded with "ostrich behavior": everything appears to be carcinogenic, so people feel helpless in the face of an overwhelming situation (6).

Given these challenges, is there anything that public health officials can do to improve the quality, usefulness, and effectiveness of the materials that they present on their Web sites? How can public health authorities attract public attention for relevant health information in a way that might lead to behavior changes?

One key is to start with good, appropriate questions: What are people interested in? What do they want from the Internet? How is it possible to assess their main interests? Is it possible to measure and track trends in public interest in health issues, and can that information be used to plan the content of a Web site? Can good Web content stand out from all the “noise” coming from the mass media? Can appropriate Web materials promote health?

Answers to some of these questions can come from qualitative research (7). Some researchers have gathered information through on-line semi-structured interviews or on-line focus groups (8). Other investigators have studied the interactions in on-line cancer support groups, where patients provide information and encouragement to each other (9, 10). Nevertheless, the usefulness of these kinds of surveys is limited by such factors as the small size of the samples and the sometimes highly biased, narrowly focused views of the participants of some on-line discussion groups.

LOG ANALYSIS OF WEB SITES

If qualitative research has those kinds of limitations, then what other choices are available to public health officials who want to improve their Web sites? One approach that has been used at the National Cancer Institute of Brazil and that we would like to recommend and to describe in this piece is “log analysis technology.”

Web log analysis was initially developed for the growing number of small Web sites that needed to evaluate their Internet “audiences.” Produced by a variety of companies, log analysis software packages provide data on Web “traffic,” such as the number of visitors to a Web site, the number of times that particular pages are viewed, and how long visitors spend at the site. Web site managers (Webmasters) use the software to analyze electronic “log files,” which list the actions that have occurred on the Web site.

With log analysis tools, it’s possible to get a good idea of how visitors have come to the Web site, if visitors choose to come back to the site, and what “paths” visitors take in navigating through the site, such as the sequence they follow in going from one Web page to another on the site.

A for-profit company, for instance, can use log analysis technology to assess the effectiveness of its on-line advertising and to learn which other Web sites have led visitors to the company’s Web site.

For their part, public health officials can use log analysis technology to better serve visitors’ needs and to more effectively communicate key public health messages—and hopefully produce behavior change. This technology enables Web designers to focus attention on the most popular pages, for up-

dates and improvements, and also indicates where to add more content. Public health institutions can better plan their health promotion strategies based on log analysis information. In this way, the Web site is used both to provide information on health issues and to generate analyses that can lead to changes in the Web site that in turn will further increase public interest in the health information.

HOW THE NATIONAL CANCER INSTITUTE OF BRAZIL HAS USED LOG ANALYSIS WITH ITS WEB SITE

An agency of the Brazilian Ministry of Health, the Brazilian National Cancer Institute (NCI) (*Instituto Nacional de Câncer*, or INCA) was created to support the country’s national health policy on cancer. The NCI was created in 1961 and reorganized in 1990, and it is responsible for cancer care delivery, cancer prevention, and cancer detection at an early stage. The NCI plans, organizes, manages, and supervises national projects and activities. The NCI and the Ministry of Health produce cancer prevention information materials for and promote events aimed at health professionals, opinion leaders, and the general public. Working with television broadcasters, the print media, and the Internet to broadly disseminate information, the NCI focuses some of its efforts around specific dates each year to spread the messages of cancer prevention and early detection. Some of these special events include World No Tobacco Day, National Fight Smoking Day, and the National Day against Cancer.

The NCI Web site (www.inca.gov.br) was established in 1997. It is the primary Web reference site on cancer for the general public in Brazil. To construct the Web indicators and reports that we describe below, we have used the WebTrends Log Analyzer software package. By helping show trends in public interest, the software assists us with our strategies, educational products, programs, projects, and activities related to health promotion and prevention. Some information disseminated by the mass media, including sensationalistic television programs and dubious Web sites, promotes questionable “alternative” cancer treatments. In order to answer the many questions raised by these media stimuli, we have used an e-mail service linked to the NCI Web site and a “frequently asked questions” feature. We have used log analysis technology since 1999 to produce data on public interest in cancer issues, and then we have applied those data along with the information coming from the e-mails to periodically reformulate the NCI educational publications and the Web site.

The information that we present below is based on an analysis of all of the log files activity for the NCI Web site from September 1999 to Au-

FIGURE 1. Trends in the number of page views per month of the “tobacco and health” page and the “how to stop smoking” page on the Web site of the National Cancer Institute of Brazil, September 1999 through August 2001

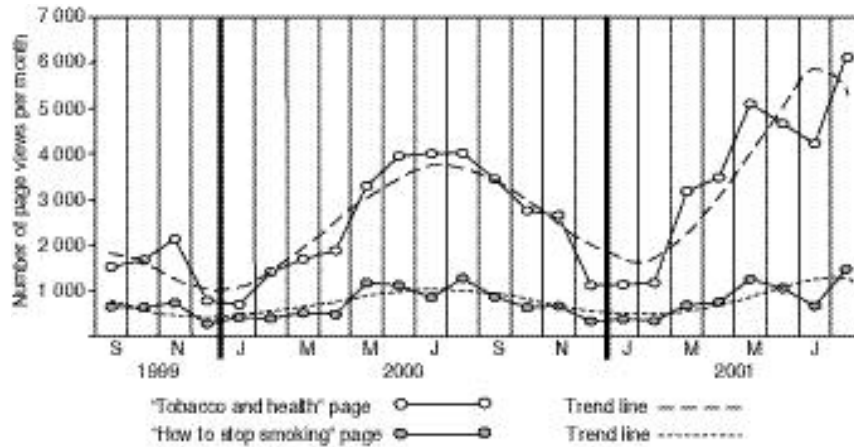
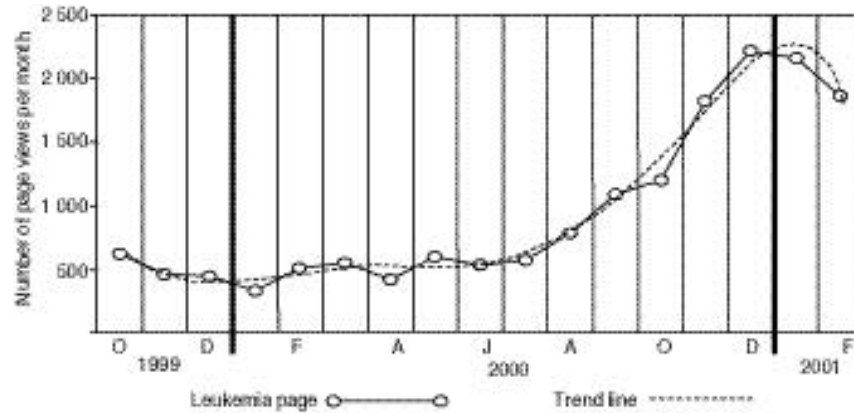


FIGURE 2. Trends in the number of page views per month for leukemia page on the Web site of the National Cancer Institute of Brazil, October 1999 through February 2001



gust 2001. In looking at the data for some of the pages on the NCI Web site, we found peaks around the time of such special days as World No Tobacco Day (May 31), National Fight Smoking Day (August 29), and the National Day Against Cancer (November 27). Figure 1 depicts the trends with two specific NCI Web pages: the “tobacco and health” page and the “how to stop smoking” page. Heavy attention from other media, including radio, television, and print outlets, may have contributed to the growth in the page views for these two Web pages.

Outside events, rather than activities that the NCI planned and carried out on its own initiative, have generated similar growth in visits to other pages of the NCI Web site. For example, Figure 2 shows the trend, from October 1999 through February 2001, in the number of views of the page on the NCI Web site that deals with leukemia. During that period a character presented in a Brazilian television soap opera was suffering from the disease.

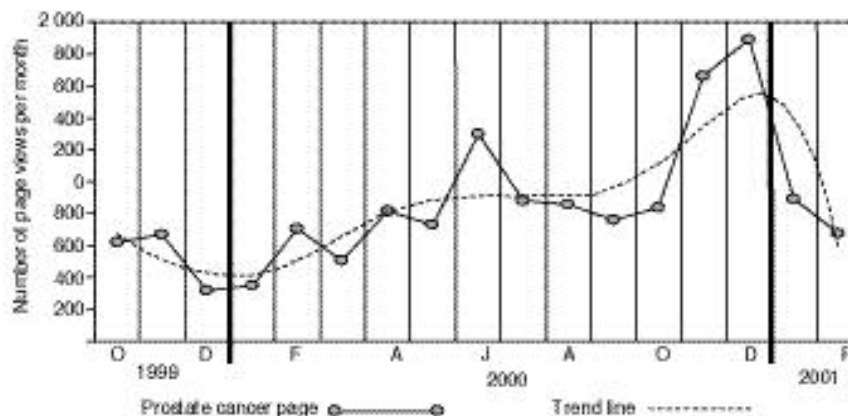
Making use of the attention that resulted from that TV program, the NCI carried out a highly successful bone marrow donation campaign, using the soap opera actress in television announcements.

Figure 3 shows the pattern, also for October 1999 through February 2001, in the number of views of the page on the NCI Web site that deals with prostate cancer. Another outside event, the illness from prostate cancer and the subsequent death of a famous Brazilian politician, the governor of the state of São Paulo, likely contributed to the growth in these page views.

CONCLUSIONS

Log analysis technology is useful for evaluating changes in public interest in health issues as shown by trends in traffic on an Internet site. As peaks occur in that traffic, adjustments can be made

FIGURE 3. Trends in the number of page views per month of prostate cancer page on the Web site of the National Cancer Institute of Brazil, October 1999 through February 2001



in the Web site to create more effective campaigns that will hopefully attract even more attention, change behavior, and improve health.

Acknowledgments. We thank Mr. Walter Zoss of the Brazilian National Cancer Institute for his invaluable help.

SINOPSIS

Evaluación mediante análisis de registros de un sitio de información de salud en Internet: la experiencia del Instituto Nacional del Cáncer, de Brasil

La informática se ha convertido en parte integrante del concepto moderno de salud pública y de las políticas sanitarias nacionales. Al propio tiempo, sin embargo, el torrente de información que se recibe a través de Internet y de otros medios de comunicación ha alejado en ocasiones la atención del público de elementos importantes que pueden ayudar a prevenir enfermedades. En este trabajo se muestra cómo mediante la tecnología de "análisis de registros" de Internet es posible evaluar y seguir las tendencias del interés del público en temas de salud. Los programas para el análisis de re-

gistros pueden ofrecer estadísticas del "tráfico" en Internet, incluyendo el número de visitas que recibe un sitio, cuáles páginas específicas de un sitio son vistas con mayor frecuencia y cuánto tiempo pasan los visitantes en el sitio. Esta tecnología puede ayudar a las instituciones sanitarias a planear sus estrategias de promoción de salud y a comunicar mensajes clave relacionados con la salud pública. Como ejemplo de cómo la tecnología de análisis de registros puede ser aplicada, se analizó la actividad del sitio en Internet del Instituto Nacional del Cáncer, de Brasil, entre septiembre de 1999 y agosto de 2001. Se observó un apreciable crecimiento del número de visitas a las páginas relacionadas con tabaco y salud y con las formas de abandonar el tabaquismo durante las campañas de lucha contra este hábito, realizadas en el marco de jornadas como el Día Mundial Sin Tabaco, el Día Nacional de Lucha Contra el Tabaquismo y el Día Nacional Contra el Cáncer. Un crecimiento considerable de las visitas a otras páginas del sitio en Internet de ese Instituto parece estar vinculado con sucesos externos, incluyendo la enfermedad y muerte de un famoso político brasileño por cáncer de próstata y una telenovela brasileña en la que un personaje padecía de leucemia. Según la experiencia del Instituto Nacional del Cáncer, de Brasil, los funcionarios públicos vinculados con la salud pueden utilizar ventajosamente la tecnología de análisis de registros para responder mejor a las necesidades de aquellos que visitan su sitio en la web, como parte de sus esfuerzos para lograr cambios en el comportamiento y mejorar la salud.

REFERENCES

- Eysenbach G. Consumer health informatics. *BMJ* 2000;320:1713-1716.
- Robertson A, Minkler M. New health promotion movement: a critical examination. *Health Educ Q* 1994;21:285-312.
- Gottlieb S. Health information on internet is often unreliable. *BMJ* 2000;321:136.
- Baillie L, Bassett-Smith J, Broughton S. Using communicative action in the primary prevention of cancer. *Health Educ Behav* 2000;27(4):442-453.
- Kreuter M. Human behaviour and cancer: forget the magic bullet. *Cancer* 1993; 72:996-1001.
- Evans R, Barer M, Marmor T, eds. Why are some people healthy and others not? The determinants of health of populations. New York: Aldine; 1994.
- Barbour RS. The role of qualitative research in broadening the 'evidence base' for clinical practice. *J Eval Clin Pract* 2000;6:155-163.
- Eysenbach G. Ethical issues in qualitative research on internet communities. *BMJ* 2001;323:1103-1105.
- Klemm P, Reppert K, Visich L. A non-traditional cancer support group. The Internet. *Comput Nurs* 1998;16:31-36.
- Sharf BF. Communicating breast cancer on-line: support and empowerment on the Internet. *Women Health* 1997;26(1): 65-84.