

RELATIVE FREQUENCY OF BREAST CANCER WEBSITE INFORMATION TOPICS: ENVIRONMENTAL RISK, PREVENTION, DETECTION, TREATMENT, AWARENESS, SOCIAL SUPPORT, AND SURVIVORSHIP

ROSE CLARK-HITT, MARLEAH DEAN, SANDI W. SMITH,
SAMANTHA NAZIONE, CHARLES K. ATKIN AND PAMELA WHITTEN

The Internet is a popular source for Americans searching for health information (Fox, 2006; Hess et al., 2005). The potential for Internet usage for breast cancer information seeking makes it important to understand the types and relative frequency of content on environmental risk factors, prevention, detection, treatment, awareness, social support, and survivorship on popular breast cancer web pages. The top search results for websites in the most popular three search engines (Google, MSN, and Yahoo) were used to identify websites and web pages, and a total of 415 web pages were analyzed. This study

Rose Clark-Hitt is a Ph.D. candidate in the Department of Communication at Michigan State University (clarkhit@msu.edu). **Marleah Dean** is an M.A. student in the Department of Communication at the University of New Mexico. Also at Michigan State University are **Sandi W. Smith**, who is a professor in communication and director of the Health and Risk Communication Center; **Samantha Nazione** who is a Ph.D. student in communication; **Charles K. Atkin**, who is Distinguished Professor of Communication and chair of communication; and **Pamela Whitten**, who is professor in the Department of Telecommunication, Information Studies, and Media, and Dean of the College of Communication Arts & Sciences. This publication was made possible by the Breast Cancer and the Environment Research Centers grant number U01 ES/CA 012771 from the National Institute of Environmental Health Sciences (NIEHS), and the National Cancer Institute (NCI), NIH, DHHS. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIEHS or NCI, NIH.

demonstrates that about half of the top appearing web pages in popular search engines provide at least some information regarding environmental risks, risk reduction behaviors, awareness, prevention, detection, and treatment, while social support and survivorship topics were less frequently appearing topics. This study used ordinal codes and individual web pages as the unit of analysis, the results of which provide greater insight than prior content analyses of breast cancer media information.

Keywords: breast cancer, breast cancer web pages, breast cancer prevention, breast cancer environmental risks

The American Cancer Society reports that 251,830 new cases of breast cancer are diagnosed in the United States every year and 40,410 die. The odds of women developing breast cancer at some time in their lives is one in eight, and one in three cases of cancer in women are breast cancer (American Cancer Society, 2008). While men may also develop breast cancer, the focus here is on women. The Internet is a valuable resource for women seeking information to help to prevent breast cancer or cope with breast cancer detection, diagnosis, treatment, or survivorship.

According to a Pew research study, 80% of Americans at one time or another have used the Internet to view information about health-related topics (Fox, 2006). While prior research has focused on the Internet as a source of information for women diagnosed with breast cancer (Fogel et al., 2002; Kirschning & Kardoff, 2008), little work has examined the Internet as a source for information regarding breast cancer prevention or the relative emphasis given to particular topics about breast cancer (Whitten et al., 2008). This work is critical because it will help to develop future websites that disseminate important information about breast cancer prevention, detection, treatment, and survivorship.

Investigators have raised doubts about the quality of breast cancer websites (Cline & Hayes, 2001). Some literature has examined the content of various media forms with regard to breast cancer (Whitten et al., 2008; Schwartz & Woloshin, 2002; Cho, 2002), however only a few studies have examined the content and design of breast cancer websites (Whitten et al., 2008; Meric et al., 2008; Meric, 2008).

Some forms of breast cancer can be preventable, thus the Internet has the potential to disseminate information women can use to personally avoid specific risks and to help others reduce their risks to the disease. The purpose of this paper is to report the types of content encountered by individuals visiting the top appearing 14 breast cancer websites in common search engine searches and accessing multiple pages of each of them. As part of Breast

Cancer and the Environment (BCERC) research project, the role of risk reduction through the avoidance of certain environmental factors is of particular interest to this study. In addition to this emphasis on prevention and environmental factors, analysis of content on awareness, detection, treatment, survivorship, and social support, and the relative emphasis that is given to each of these topics is provided here. In sum, this research is important because it provides a comprehensive review of web based breast cancer information topics that pertain to women in all phases of dealing with breast cancer, from preventing it to detecting and treating it, and as a result this content analysis may provide areas for improvement of such websites.

BREAST CANCER INTERNET USE AND BREAST CANCER CONTENT ANALYSES

Women have a greater likelihood than men of searching for health information in general on the Internet (Fox, 2006). Furthermore, research demonstrates that many women living with breast cancer use breast cancer websites for information about the disease; in one study, 40% of the breast cancer patients sampled indicated that they used the Internet to learn about breast health issues (Fogel et al., 2002). Kirschning and Kardoff (2008) reported on Internet use among women at all stages of breast cancer, from diagnosis to rehabilitation to homecare. The proportions of women citing various reasons for using the internet with regard to breast cancer, included: being in better control of the disease (54%), obtaining information in understandable language (58%), the desire to utilize all available information sources (94%), the desire for a good basis of knowledge about the disease (74%), and a feeling of possessing better control over breast cancer (54%) (Kirschning & Kardoff, 2008). Clearly, the web is a preferred source of information about breast cancer for women, however searching for information about environmental causes of breast cancer and prevention techniques was not noted.

Cline and Hayes (2001) argued for the importance of future research regarding Internet issues such as design quality, quality of information provided, and accessibility of information. Several studies have heeded their call and examined the content of coverage of breast cancer related topics in the media with an emphasis on prevention-related topics. Atkin et al. (2008) examined the nature of breast cancer coverage in newspapers, news magazines, and television news broadcasts. This analysis examined content regarding environmental risks and prevention behaviors, among other topics. About a third of stories focused on preventing breast cancer, and most of these prevention topics focused on exposure to certain pharmaceuticals. Few news stories, however, discussed risk-reducing

lifestyle behaviors such as avoiding environmental chemicals, addressing obesity, and promoting exercise. Similarly, Cho (2006) examined 602 news stories from major television networks from 1974 to 2003, finding that the topics of prevention and treatment increased during that time period. Specifically, the topics of prevention and treatment increased in number and the percentage of total stories across time periods, with prevention moving from 2% to 5% from the first to second and from 5% to 9% of stories in the second to third periods. Treatment moved from 9% to 10% in the first to the second time period and from 10% to 18% in the second to third time periods (Cho, 2006). Schwartz and Woloshin (2002) focused exclusively on news coverage pertaining to mammography and the drug Tamoxifen as a prevention technique; they found that most articles promoted mammography, and most encouraged people to use caution when considering using Tamoxifen. Finally, Brown et al. (2001) examined news about the potential environmental causation of breast cancer, finding little coverage of environmental causation.

While a number of content analyses have examined the breast cancer information in television and newspaper sources, few studies have measured the content of websites. Meric et al. (2008) examined the content of the top 200 most popular breast cancer websites for type of content as well as quality, with quality determined by whether websites cited authorship, current information, and reference. The more popular sites were more likely to feature medical facts, opportunities for psychosocial adjustment, human-interest stories, ongoing trials, message boards, chat sites, place for questions, and results of clinical trials; there was no difference, however, between more and less popular websites regarding quality of information in terms of whether items were displayed including authorship, source of information, dates of updates, advertising policies, and any conflict of interest. Furthermore, Whitten et al. (2008) examined the content of 157 breast cancer websites in terms of theoretical components of common health communication theories, as well as the quality of website design. This study examined the presence of theoretical components from the Theory of Planned Behavior (Ajzen, 1985), the Transtheoretical Model (Prochaska, 1984), and the Extended Parallel Process Model (Witte, 1992), finding little use of behavior change theoretical components in these websites. According to this study, the majority of websites incorporated desirable design features such as listing sources of information, functional task bars, accurate information, and including relevant graphics. These recent research endeavors provided a focus on website features and theories, but little emphasis was given to amount of content. The content that was assessed was determined in terms of whether it was present or absent rather than assessing its relative frequency. Relative frequency of content in media can suggest the relative importance of a given topic as evidenced in agenda setting theory research (Jones, Denham, & Springston, 2006) and can lead to proactive health behaviors such as breast cancer screening.

BREAST CANCER RESEARCH

The website content on breast cancer of interest in this study includes the following topics: environmental risk, prevention, awareness, detection, treatment, social support, and survivorship. Environmental factors are of particular importance in light of recent research finding that there is little coverage of these topics (Atkin et al., 2008; Cho, 2006; Woloshin, 2002; Brown, 2001) despite increasing evidence of environmental risks (BCERC, 2008).

Smith et al. (2009) investigated memorable messages with regard to breast cancer, finding detection messages were the most frequent types of memorable messages (37%), followed by awareness (31%), and treatment (26%), with just 6% of memorable messages relating to prevention. In addition, the two related topics of survivorship and supportive communication are of interest in this study. Kreps (2008), noted that "Cancer survivors have unique communication needs to help them cope with the fear of their cancer recurring. Survivors also need to access social support... to enable them to readjust to their everyday lives" (p. 165). People are defined as breast cancer survivors from the moment of diagnosis onward (Thompson & O'Hair, 2008). Kreps (2003) argues that survivorship is part of the continuum of cancer care in addition to prevention, detection, diagnosis, treatment, and end of life care, and that social support can improve cancer survivorship. Survivorship includes issues such as physical and psychological adjustment, coping with side effects of treatment such as hair loss, and fear of recurrence.

Several studies have examined the role of social support among women with breast cancer. Katapodi et al. (2002) reported that women with greater social support had greater adherence to breast cancer screening guidelines. Furthermore, Winselberg et al. (2003) reported an evaluation of a mediated Internet support group, "Bosom Buddies", for women with breast cancer with regard to improvements in coping with psychological distress. The study found that the women assigned to a social support group condition over 12 weeks had reduced scores on measures of depression, stress, and trauma.

RESEARCH QUESTION

The goal of this research is to examine the relative amount of content related to six dimensions of information present in the top breast cancer websites in order to answer the following research question.

RQ: What is the extent of coverage on the top appearing websites for the topics of prevention, environmental risks, detection, treatment, social support, and survivorship?

These topics are of importance to women's health outcomes because this type of content on websites may help women to prevent breast cancer, as well as make informed decisions with their health care providers for detection and treatment. Additionally, support and survivorship are important for women who are dealing or have dealt with breast cancer.

METHOD

This study proceeded in several stages. First, the most commonly viewed breast cancer websites were identified. Second, a coding scheme was employed to evaluate the relative amount of content on the websites in terms of environmental risk, detection, treatment, awareness, prevention, survivorship, and social support.

Web Pages and Unit of Analysis

Breast cancer" and "breast cancer and environment" were entered as search terms on the three most commonly used search engines (Google, MSN and Yahoo), based on Nielsen/NetRatings (Bausch, 2007). The first 30 websites from each of these three search engines with each of the two search terms above were recorded, providing a total of 180 websites. The search result rank order was listed in a column next to each web page. The sort function in Excel was used to match up identical domain names to determine overlap in results. Because the goal of the study was to examine the population of the top 14 websites and the web pages within one click of the home page on each (N=415 pages), all returns were classified based on the first level domain name (e.g., www.breastcancer.org/symptoms and www.breastcancer.org/treatment were classified in the www.breastcancer.org category), as suggested by Weare and Ying Lin (2000). The final list was derived by using the top 14 sites based on the search result rank order. Criteria were applied for inclusion in the list of websites to be analyzed. First, websites that were paid advertisements were not included. Second, news source websites were not included (e.g., "Yahoo News"). Third, blogs and chat room web pages within websites were not included because of the continually changing content of those pages and because the content is not necessarily controlled by the organization running the site. The websites identified for this research are listed in Table 1.

Weare and Ying Lin (2000) argued that the individual web page is a strong choice for unit of analysis as opposed to an entire web site. The advantage of using individual web pages as the unit of analysis is that it provides a comprehensive perspective of the specific content on each page, as there can be great variability in content from page to page within a website. Most of the websites chosen for this sample have hundreds of linked pages within

Table 1. Web Addresses and Organizations for Websites in Sample	
Organization	Web Address
Cornell University Program on Breast Cancer and Environmental Risk Factors	envirocancer.cornell.edu
Breast Cancer and the Environment Research Centers	bcerc.org/home.htm
BreastCancer.Org	breastcancer.org
Breast Cancer Fund	breastcancerfund.org
National Breast Cancer Coalition	stopbreastcancer.org
National Institute of Environmental and Health Sciences	niehs.nih.gov/research/supported/centers/breast-cancer/index.cfm
Yahoo Health	health.yahoo.com/breastcancer-overview/breast-cancer-topic-overview/healthwise-tv3617.html
Mayo Clinic	mayoclinic.com/health/breast-cancer/DS00328
Oncolink	oncolink.org/types/article.cfm?c=3&s=5&ss=33&id=8320
The National Cancer Institute	cancer.gov/cancerinfo/types/breast
The National Cancer Institute	cancer.gov/cancerinfo/pdqatmbreas/helathprofessional/allpages/print
Wikipedia	en.wikipedia.org/wiki/Breast_cancer
BreastCancer.net	breastcancernet.org
American Cancer Society	cancer.org/docroot/CRI/CRI_2x.asp?sitearea=LRN&dt=5

the website. To balance evaluation of the content in each website with the overwhelming number of pages on each site, only pages that could be reached within one click of the home page were analyzed (Weare & Ying Lin, 2000) for a total of 415 web pages analyzed. This method provides a comprehensive view of the type of variability among web pages within a single website. The content, or snapshots of the websites, was captured by PDF documents for ease of coding. All searches were conducted in the USA between late October, 2008 and June, 2009.

Measures

For each category of interest (awareness, environmental risks, prevention, detection, treatment, survivorship, and support) several codes were specified. Each is discussed below in its respective category.

Awareness. The topic of awareness is defined here as information about prevalence of breast cancer and scientific information explaining breast cancer. Prevalence is an indicator of the overall level of risk of disease (Atkin et al., 2008). It conveys a level of threat about how likely it is that an individual may contract a disease. Prevalence may be expressed through odds, rates of occurrence, numbers of people with a disease, as well as descriptors such as “widespread” or “increasing,” thus the variables for prevalence are: odds of developing breast cancer, number of people with breast cancer, and verbal descriptors of prevalence such as “widespread” or “increasing.” Scientific information about breast cancer includes diagrams, information about the stages of the disease, and types of breast cancer.

Environmental Risks. Different researchers provide varying levels of differentiation within the category of “environmental risks.” Atkin et al. (2008) differentiated environmental risks as primary risk and secondary, with primary risks including family heredity, history of breast cancer, breast density, and presence of BRCA1 and BRCA2 genes. Secondary environmental risks are those for which avoidance is prudent, including environmental contaminants, chemical contaminants, exposure to cigarette smoke, personal use of tobacco products, exposure to certain hormones (e.g., hormone replacement therapy), use of pharmaceuticals, lack of participation in physical exercise, consuming unhealthy foods, and obesity.

Prevention. Prevention has three dimensions: individual, parental, and collective behaviors. Prevention is closely linked with the environmental risks category. Individual prevention behaviors include specific information about behaviors individuals may engage in to prevent breast cancer, including eating a healthy diet, avoiding tobacco use, exercise,

moderate or no alcohol consumption, and use of drugs such as Tamoxifen. Parental prevention behaviors include encouraging daughters to engage in particular behaviors such as healthy eating, exercise and avoiding environmental contaminants. Collective prevention behaviors involve getting involved with policy actions such as "...supporting policies to restrict local chemical emissions" (Atkin et al., 2008, p. 5). Furthermore, prior research (Atkin et al., 2008) has also included donations as a feature of preventive behavior because these activities are promoted as actions to prevent breast cancer, thus donations were coded as prevention in this study.

Detection. Detection involves actions to find breast cancer. Detection may be performed by trained medical professionals or by individual women performing breast self exam (BSE). Detection performed by medical professionals includes mammography, clinical breast exams, and magnetic resonance imaging (MRI) (Atkin et al., 2008). Symptoms such as breast lumps are also involved in detection.

Treatment. Treatment involves the eradication of breast cancer from individuals. Treatment is supervised or directed by health care professionals. Attempts to treat breast cancer coded in this study are chemotherapy, drugs, surgery, radiation, hormone replacement therapy, and holistic or alternative medicine.

Survivorship. Survivorship codes included fear of recurrence of breast cancer among survivors, healing, post-treatment side effects such as hair loss, and physical/psychological adjustments (Kreps, 2003). Physical and social/psychological adjustments included topics such as having to ask for help lifting heavy items, intimacy with partner, and talking to coworkers or employers.

Social Support. Social support was defined in terms from Cutrona and Suhr (1992): information support (e.g. advice, factual support), tangible support (e.g., offers for assistance with resources such as helping cook meals or picking up children from school), emotional support (expressing caring or empathy), esteem support (expressing respect or confidence person's competence), and social network support ("communicating with a group of similar others") (Cutrona & Suhr, 1992). Thus, the categories of social support are information, tangible, emotional, esteem, and network support.

Coding

An ordinal coding scheme was employed here. Ordinal coding provides an advantage over most prior literature employing nominal coding schemes; rather than simply coding for the presence or absence of a concept, codes were for whether the topic was the major emphasis (overall theme or subject of individual webpage to approximately one quarter of the content in the page), minor emphasis (i.e., link, tab name, or a brief mention of the topic), or no content (no information provided on category) for every topic that was coded. An ordinal coding scheme is advantageous because it provides a more detailed picture of the content devoted to particular topics.

Coder Training and Inter-Coder Reliability

For training purposes, two authors practiced coding websites together on websites not included in the sample. They then established reliability on just over 10% of the total sample. The inter-coder reliability values are reported for each variable in tables 2 through 8 below the frequencies. The SPSS macro for Krippendorff's alpha was used to calculate the reliability. Krippendorff's alpha is a statistic that is suggested for use with ordinal level data (Hayes & Krippendorff, 2007), as other reliability statistics such as Scott's Pi and Cohen's Kappa are limited to use with nominal data only (Hayes & Krippendorff, 2007) and are thus may not used here because this data set is ordinal. For some variables the Krippendorff's alpha statistic was not applicable due to a lack of covariation in the data (e.g., nearly all applied codes for a particular variable were "0"), therefore percentage agreement values are reported in those cases.

RESULTS

Frequencies are reported for each variable below in the categories of environmental risks, prevention, detection, treatment, social support, and survivorship. These frequencies provide information about the extent of information available to the audiences of these web pages. Tables 2 through 8 include the Krippendorff's alpha inter-coder reliability coefficient (Krippendorff's alpha values ranged from 73% to 100%), and percentage agreement, where applicable (ranging from 81% to 100%), for each variable reported.

Table 2. Environmental Risks

content	Overall Environmental Risk	Hormones	Pharmaceuticals	Exposure to chemicals	Exposure to pesticides	Exposure to smoke	Personal Tobacco use	Lack of Exercise	Unhealthy diet	Obesity
No Content	50.4	87.6	99.8	87.8	96.7	96.3	96.2	95.2	88.8	95.5
Minor	36.5	10.5	0.1	7.2	2.6	1.4	3.6	4.8	10.7	4.3
Major	13.1	2	0	5	0.7	0.2	0.2	0	0.4	0.1
K alpha	.8	.88*	1.0*	.92*	.81	1.0*	.79	.94*	.90	.79

Environmental Risks

Close to half of web pages included at least a mention of some sort of environmental risk with regard to breast cancer, and for 13% of them environmental risks were a major emphasis. For hormones, about 10% had a minor emphasis and 2% had a major emphasis. For eating an unhealthy diet, about 11% of web pages had a minor emphasis and 1% had a major emphasis, while for exposure to chemicals, about 7% had a minor emphasis, and 5.1% had a major emphasis. Both obesity and lack of exercise were mentioned in about 4% of websites. Use of pharmaceuticals, exposure to pesticides, exposure to smoke, and personal use of tobacco were each mentioned by fewer than 4% of websites. Please see Table 2 for complete results.

Awareness

Awareness was a topic on about 52% of the websites that were coded with 16% presenting it as a major topic. Nearly 13% of the web pages had a major emphasis on

Table 3. Awareness

content	Overall Awareness	Odds	Number of deaths	Total number of cases	Rate of breast cancer	Rate phrases	Prevalence Comparison	Scientific information
No Content	47.3	92.6	97.1	95.9	99.5	100	99.3	59.7
Minor	36.5	6.7	2.6	4.1	0.5	0	0.7	27.9
Major	16.2	0.7	0.2	0	0	0	0	12.4
K alpha	.79	.79	.96*	.92*	.98*	1.0*	1.0*	.81*

Table 4. Prevention

content	Overall Prevention	Individual Prevention	Parental Prevention	Collective Prevention	Donating Time or Money
No Content	39.1	86.2	98.8	84.7	74.7
Minor	45.1	9.5	1.2	6	23.6
Major	15.7	4.3	0	9.3	1.6
K alpha	.79	.73	.98*	.98*	.95

scientific information, with about 28% having a minor emphasis, or mention, of scientific information. The most common awareness topic was scientific information about breast cancer, while prevalence of the disease received less emphasis across web pages. About 4% of web pages mentioned the number of cases of breast cancer, with about 3% mentioning number of deaths from the disease. Close to 6% of the web pages mentioned the probability (or odds) of developing breast cancer, about .5% of sites having a major emphasis on odds of developing breast cancer. Table 3 presents the complete results for Awareness.

Prevention

Prevention was a topic on about 61% of the websites that were coded with about 16% presenting it as a major topic. Within the category of prevention, about 4% of websites had a major emphasis and about 9% had a minor emphasis on individual prevention efforts, parental prevention had the least coverage of any prevention category with only a mention

Table 5. Detection

content	Overall Detection	Breast Self Exam	Mammogram	Clinical Breast Exam	X-Ray	MRI	Symptoms
No Content	47.5	94.7	83.1	95	94.3	93.8	68.2
Minor	43	5	14.3	5	5.7	5.7	31.3
Major	9.6	0.2	2.6	0	0	0.5	0.4
K alpha	.88	.79	.96	.98*	.98*	1.0*	.79

on about 1% of the web pages, whereas collective prevention was the major emphasis of about 9% and minor emphasis of about 6% of web pages in the sample. Nearly a quarter of web pages had a minor emphasis on donating time or money in efforts to help prevent breast cancer, with about 1.5% having a major emphasis on donating time or money. Please see Table 4 for complete Prevention results.

Detection

Detection was a topic on about 52% of the websites that were coded with close to 10% presenting it as a major topic. The most commonly occurring detection procedure was mammograms with websites nearly 15% having a minor emphasis on mammograms and nearly 3% having a major emphasis. Breast self-exams and clinical breast exams were each mentioned in about 5%, and x-rays and MRI were each mentioned in about 6% of the web pages coded here. Finally, symptoms were the most commonly appearing detection topic, with nearly a third of web pages mentioning them. Table 5 presents the complete Detection results.

Treatment

Treatment was a topic on about 62% of the websites that were coded with over 21% presenting it as a major topic. The most common treatment content concerned chemotherapy, with nearly 30% of web pages coded here having a minor emphasis and nearly 5% having a major emphasis on it. Radiation was a minor emphasis in about 17% and major emphasis in about 3% of pages, while surgical procedures were a minor emphasis in about 19% and a major emphasis in about 5% of web pages. Holistic medicine was the minor emphasis in about 4% and in less than 1% of web pages. Please see Table 6 for complete Treatment results.

Survivorship

Survivorship was a topic on about 39% of the websites that were coded with only 0.5% presenting it as a major topic. The most common survivorship topic was side effects such as hair loss and nausea, with about 5% web pages including such content. Each of the other survivorship topics was only the minor emphasis or had a major emphasis in less than a combined 3% of web pages. Table 7 presents the breakdown of Survivorship categories.

Table 6. Treatment

content	Overall Treatment	Chemotherapy	Radiation	Surgical Procedures	Holistic Medicine
No content	37.7	66.3	79.7	76.6	95.2
Minor Mention	40.8	29.1	17.4	18.9	4.3
Major	21.5	4.5	2.8	4.5	0.4
K alpha	.93	.94	.93	.9	.85

Social Support

Social support was a topic on about 22% of the websites that were coded with only 2.2% presenting it as a major topic. Most prevalent was network support with minor emphasis on about 13% web pages, typically with regard to links for chat rooms or online support groups. Emotional, esteem, instrumental, and informational support were each a minor emphasis in four or fewer of web pages, with no web pages having a major emphasis on these topics. Please see Table 8 for a breakdown of Social Support coding.

DISCUSSION

Principal Results

Analysis of the content of the top appearing web pages in popular search engine provides insight in to the types of information that individuals searching for information regarding breast cancer and the role of environmental factors in breast cancer typically encounter. This information has implications for website and health campaign designers as well as practitioners in terms of addressing areas in which there is little information available to Internet users. In particular, the relative lack of information on survivorship and social support is an area that needs to be remedied. While some may argue that survivorship and social support are not necessary topics on breast cancer web pages because women could enter different search terms to locate such resources, it may be beneficial for general breast cancer websites to at least direct women to such resources if needed, even if those resources are on different websites.

Table 7. Survivorship

content	Overall Survivorship	Healing	Fear of Recurrence	Side effects of treatment	Post treatment adjustment
No Content	61.3	99.8	97.4	93.8	96
Minor	34.8	0.2	2.6	4.8	3.5
Major	0.5	0	0	1.4	0.4
K alpha	.72	.98*	1.0*	.82	.86

*Indicates percent agreement. In several cases Krippendorff's alpha could not be calculated because of lack of covariance in the data (e.g., nearly all applied codes for a particular variable were "0").

Environmental risks were mentioned on about half of the websites; this is not surprising however, given that one of the search terms was "breast cancer and environment." The most common environmental risk topic was the role of hormones, specifically hormone replacement therapy or endocrine disruption in which environmental chemicals mimic estrogen, thereby increasing cancer risk. The second most common topics were chemical exposures in the environment and the role of eating a healthy diet. However, very few web pages included content pertaining to the role of obesity, pesticide exposure, pharmaceuticals, and exposure to smoke and tobacco.

The most common awareness topic was scientific information regarding breast cancer, including items such as diagrams of breast cancer, information about the types of breast cancer, and stages of breast cancer. Over 40% of web pages had at least a mention of scientific information explaining breast cancer, while far fewer websites included information about the probability of developing breast cancer and the prevalence of breast cancer.

The majority of prevention information was targeted at individuals, with virtually no content pertaining to parental behaviors to help prevent breast cancer among daughters or collective activities by communities to mitigate community breast cancer risks. This is an interesting finding particularly in light of the importance of parental protection for daughters in terms of avoiding particular environmental factors in formative years, an emphasis of the breast cancer and the environment research project (Atkin et al., 2008). The relative

frequency of prevention as a topic was high in comparison to the mere 6% of memorable messages about breast cancer that were on this topic. The fact that almost 61% of the web pages made some mention of prevention means that the topic should be seen as important by the public and get discussed in media and interpersonal channels. These discussions could then lead to future memorable messages and ultimately to more frequent prevention behavior.

Most detection information on these web pages was general in nature followed by information on mammograms and to a lesser extent clinical breast exams, x-rays, and MRI procedures. Breast self examinations were not frequently mentioned on the web pages.

Treatment information that was general was most frequent with chemotherapy the most frequent specific treatment option mentioned. It was followed by radiology and surgery. Holistic approaches were infrequently mentioned.

As noted earlier, the relative frequency of health topics, such as the various aspects of breast cancer reviewed herein, can be viewed through an Agenda Setting perspective. Relatively high amounts of content in a medium, such as the web, can create beliefs in information seekers that a topic is important. Therefore, the assessment of relative frequency of health topics in a mass medium, such as the web, is critical to understanding beliefs and behaviors that ensue as a function of consumption of that medium.

Websites provide varying degrees of information and content on various breast cancer topics. While it is not possible to make definitive comparisons between websites, particular websites included relatively high proportions of content in the areas of awareness, environmental risks, prevention, detection, treatment, survivorship, and social support. Those websites with the greatest percentage of *awareness* information across web pages were Wikipedia, Yahoo, and Cornell. With regard to *environmental risks*, bcerc.org (the Breast Cancer and the Environment Research Centers), breastcancer.org, and breastcancerfund.org, had the highest percentage of web pages including at least a mention of environmental risk factors. With regard to *prevention*, the breastcancerfund.org, stopbreastcancer.org, and breastcancer.org websites had the highest percentages of web pages including at least a mention of prevention. Websites with the greatest percentages of *detection* content were breastcancer.org, the Mayo Clinic, and Yahoo websites. The BCERC, Cornell, and Cancer.gov websites had the highest percentage of web pages with at least a mention of *treatment*. Finally, Cornell and Yahoo had the most with *survivorship* topics, and breastcancer.org and yahoo had the most *social support* content.

Comparison with Prior Work

The present study differs from prior work examining breast cancer information in several regards. First, while prior work has examined the content of coverage of prevention and environmental factors in the news media, to our knowledge, none have specifically examined the content of websites with regard to these topics. Due to the large number of people seeking breast cancer information on the Internet, and the availability of scientific information about preventing breast cancer and environmental factors, it is particularly important to understand the type of content available on these topics to have a sense of how much information about prevention is available to this audience. Second, this study is unique from other breast cancer content website analysis studies because it examines individual web pages as the unit of analysis rather than entire websites, providing more specific information regarding individual pages within websites, thus providing a more comprehensive view of how such information is distributed within websites. Third, instead of using dichotomous coding for presence or absence of particular content, ordinal scaling was used to provide a more detailed description of the content available to Internet users when they search for breast cancer information. Having such nuanced information is advantageous because it provides a more comprehensive set of information about the relative proportions of content available to Internet Users, for example, whether the page merely mentioned a topic or focused on the topic in detail on the page.

Future Directions

While the purpose of this study was to provide a largely descriptive analysis, of content available on websites, this study does provide some future directions for theoretical study of this topic. For example, due to the potential risks for breast cancer involved with environmental factors, it may be useful to use the extended parallel process model (Witte, 1992) to examine the degree to which websites addressing environmental risks provide messages with a threat demonstrating risk in terms of severity and susceptibility, accompanied by messages describing efficacy (self efficacy and response efficacy) for individuals to engage in behaviors to prevent breast cancer.

In addition, a new study could follow the lead of Jones, Denham, and Springston (2006) and use an agenda setting lens to determine exposure to web content on breast cancer and how it relates to importance of health topics and proactive health behaviors.

Designers of breast cancer web sites might also profit from using the information gleaned here to check the content of their web pages. In particular, survivorship and social support aspects of the content might need to be enhanced.

Limitations

Among the key limitations of this study, coding was confined to those web pages within one click of the main page rather than encompassing a broader array of pages. Information on secondary sources of risks might have been found if multiple links were pursued, although the major types of content are likely to be included with the procedure used here. Second, of the hundreds of breast cancer websites appearing on the web, fourteen major websites were analyzed. Although this is a small proportion of all websites, the selection process yielded sites that generated the most traffic. Third, because the study used the search terms “breast cancer and environment” as well as “breast cancer,” the content analysis likely elevated the prevalence of environmental risk factors. This may not be representative of what women would find on the Internet if they sought information about general breast cancer risks.

Conclusion

This study examined the content of breast cancer web pages in terms of environmental risks, awareness, prevention, detection, treatment, survivorship, and social support. This study contributes to the literature because while prior studies have examined media sources for such content, no study has examined web pages. Further, the ordinal coding scheme used for this analysis provides a more nuanced view of the proportion of content in each of these areas provided by each website beyond the coding scheme used in prior content analytic work on breast cancer in media sources.

REFERENCES

- American Cancer Society (2009). Breast Cancer Facts & Figures. Retrieved October 21, 2009 from http://www.cancer.org/downloads/STT/F861009_final%209-08-09.pdf.
- Atkin, C., Smith, S., McFeters, C., & Ferguson, V. (2008). A comprehensive analysis of breast cancer news coverage in leading media outlets focusing on environmental risk and prevention. *Journal of Health Communication, 13*, 3-19.
- Azjen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl and J. Beckman (Eds.), *Action control: From cognition to behavior* (pp. 11-39). New York: Springer.
- Bausch, S. (2007). Nielsen/Netratings announces May U.S. search share rankings. Retrieved October 10, 2007 from http://www.nielsen-online.com/pr/pr_070620.pdf.

- Breast Cancer and the Environment Research Centers (BCERC) Research Retrieved October, 15, 2008 from <http://www.bcerc.org/research.htm>.
- Brown, P., Zavestoki, McCormick, S., Mandelbaum, & Luebke, T. (2001). Print media coverage of environmental causation of breast cancer. *Sociology of Health and Illness*, 23, 747-775.
- Cho, S. (2006). Network news coverage of breast cancer, 1974 to 2003. *Journalism and Mass Communication Quarterly*, 83, 116-130.
- Cline, R. J. & Haynes, K. M. (2001). Consumer health information seeking on the Internet: the state of the art. *Health Education Research*, 16, 671-692.
- Fogel, J., Albert, S., Schnabel, F., Ditkoff, B., & Neugut, A. (2002). Use of the Internet by women with breast cancer. *Health Psychology*, 21, 398-404.
- Fox, S. (2006). Seeking health online. PEW internet and American life project. Retrieved August 2, 2009 from <http://pewresearch.org/pubs/265/seeking-health-online>.
- Hayes, A. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77-89.
- Jones, K. O., Denham, B. E., & Springston, J. K. (2006). Effects of mass and interpersonal communication on breast cancer screening: Advancing agenda-setting theory in health contexts. *Journal of Applied Communication Research*, 34, 94-113.
- Katapodi, M., Facione, N. C., Miaskowski, C., Dodd, M. J., & Waters, C. (2002). The influence of social support on breast cancer screening in a multicultural community sample. *Oncology Nursing Forum*, 29, 845-852.
- Kirschning, S., & Kardorff, E. (2008). The use of the Internet by women with breast cancer and men with prostate cancer-results of online research. *Journal of Public Health*, 16, 133-143.
- Kreps, G. (2003). Communication, cancer risk, and quality of life. *Health Communication*, 15, 161-169.
- Meric, F., Bernstram, N., Mirza, N., et al. (2002). Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. *British Medical Journal*, 324, 577-581.
- O'Hair, D., Villagren, M., Wittenberg, E., Brown, K., Ferguson, M., Hall, H., & Doty, T. (2003). Cancer survivorship and agency model: Implications for patient choice, decision making, and influence. *Health Communication*, 15, 193-202.
- Prochaska, J.O. (1984). *Systems of psychotherapy: A transtheoretical analysis* (2nd Ed.). California: Brooks-Cole.
- Schwartz, L.M., & Woloshin, S. (2002). New media coverage of screening mammography for women in their 40s and tamoxifen for primary prevention of breast cancer. *Journal of American Medical Association*, 287(23), 3136-3142.
- Smith, S. W., Nazione, S., LaPlante, C., Kotowski, M., Atkin, C., Skubisz, C., & Stohl, C. (2009). Topics and sources of memorable breast cancer messages and their impact on prevention and detection behaviors. *Journal of Health Communication*, 14, 293-307.
- Thompson, S., & O'Hair, D. (2008). Advice-giving and the management of uncertainty for cancer survivors. *Health Communication*, 23, 340-348.

- Weare, C., & Lin, W.Y. (2000) Content analysis of the world wide web: Opportunities and challenges. *Social Science Computer Review*, 18, 272-292.
- Whitten, P., Smith, S., Munday, S., & LePlante, C. (2008). Communication assessment of the most frequented breast cancer websites: Evaluation of design and theoretical criteria. *Journal of Computer Mediated Communication*, 13, 880-911.
- Winselberg, A., Classen, C., Alpers, G., Roberts, H., Koopman, C., Adams, R., Ernst, H., et al. (2003). Evaluation of an Internet support group for women with primary breast cancer. *Cancer*, 97, 1164-1173.
- Witte, K. (1992). Putting the fear back in fear appeals: The extended parallel process model. *Communication Monographs*, 59, 329-363.