Journal of Behavioral Medicine, Vol. 27, No. 5, October 2004 (© 2004)

Use of the Internet for Information and Support: Disclosure Among Persons With Breast and Prostate Cancer

Jason E. Owen,[1,5] Joshua C. Klapow,[2,3] David L. Roth,[4] and Diane C. Tucker[2]

Accepted for publication: February 5, 2004

The present study examined the feasibility of evaluating online communication of cancer patients using an automated content analysis program modified for application to cancer-related communication. Public messages posted to the Breast Cancer Discussion List and the Prostate Problems Mailing List were content analyzed using an augmented version of Linguistic Inquiry and Word Count to evaluate communication styles within these two cancer types. Breast cancer patients were more likely to submit multiple messages to the list and made greater use of words related to emotional disclosure and cognitive processing compared with prostate cancer patients. Prostate cancer patients were less likely to seek emotional support or repeated interaction with other patients, and more of their communication focused on cancer-related information. Use of cancer-specific word libraries significantly increased word identification within these samples. Content analysis of online communication

[1]Division of Cancer Prevention and Control Research, School of Public Health and Jonsson Comprehensive Cancer Center, University of California, Los Angeles, California.
[2]Department of Psychology, The University of Alabama at Birmingham, Birmingham, Alabama.
[3]Department of Health Care Organization and Policy, The University of Alabama at Birmingham, Birmingham, Alabama.
[4]Department of Biostatistics, The University of Alabama at Birmingham, Birmingham, Alabama.
[5]To whom correspondence should be addressed at UCLA Division of Cancer Prevention and Control Research, School of Public Health and Jonsson Comprehensive Cancer Center, 650 Charles Young Drive South, A2-125 CHS, Box 956900, Los Angeles, California 90095-6900, e-mail: jeowen@ucla.edu.

0160-7715/04/1000-0491/0

appears to be a promising method for detecting communication differences among subgroups of cancer patients.

KEY WORDS: cancer; Internet; communication; support.

INTRODUCTION

A rapidly growing number of patients with cancer are accessing the Internet to obtain support from others and information about their illness (Fogel *et al.*, 2002; Meric *et al.*, 2002). However, little is known about how patients utilize these resources to cope with the cancer experience, in part because the medium of the Internet is not conducive to performing controlled studies (Epstein and Klinkenberg, 2001). Specifically, samples of individuals using the Internet are unlikely to represent a random sample from a population, and the relative anonymity of Internet-based interactions severely limits the ability to control for possible effects of sociodemographic or psychological characteristics. Given these limitations, it may be possible to quantify the ways in which such groups are used by men and women with cancer by analyzing the communication that takes place in Internet discussion groups. Characterizing patterns of behavior exhibited by an increasingly large subset of patients who are actively using the Internet will improve our understanding of how online services and communities are used to cope with diagnosis, treatment, and survival.

In a qualitative content analysis of online discussion groups, Klemm *et al.* (1999) have shown that men with prostate cancer are more likely to seek and provide information, whereas women with breast cancer are more likely to share personal experiences and provide support online (Klemm *et al.*, 1999). The Klemm *et al.* (1999) findings suggest that gender differences in communicating the cancer experience are not unlike gender differences in communicating about other distress-inducing events. In an analysis of 13 studies of the written self-disclosure paradigm developed by Pennebaker and colleagues, Smyth (1998) has shown that men exhibit greater benefit than women from written emotional expression. This effect may be partially explained by findings that men generally exhibit lower levels of emotional expressivity than do women (Brody, 1996; Campbell, 1996; Polce *et al.*, 1998 ; Sells and Martin, 2001). In studies of the effects of writing about emotionally-charged events, positive health outcomes have been tied to three specific linguistic patterns: 1) high use of positive emotion words (e.g. happy, joyful), 2) moderate use of negative emotion words (e.g. sad, angry), and 3) increasing use over time of words related to insight and causation of the event (e.g. because, therefore; Esterling *et al.*, 1999; Pennebaker, 2000). It has been hypothesized that these linguistic patterns may influence health by allowing individuals to integrate disparate affective and cognitive

attributions of distress-inducing events into an organized and coherent narrative (Pennebaker, 1997, 2000). Several studies have suggested that emotional expression is associated with lower health-care resource utilization, higher levels of overall health, and lower distress over time among breast cancer patients (Classen *et al.*, 1996; Stanton *et al.*, 2000).

Internet discussion groups for cancer are similar to expressive writing tasks in that participants write about the distress-inducing event of cancer, potentially sharing thoughts and feelings that are perceived to be too difficult to share with loved ones (Sharf, 1997). Unlike expressive writing tasks, however, participants in Internet discussion groups are offered little or no guidance regarding their communication and are not journaling privately but to a public forum, often with the explicit goal of receiving feedback, support, and information from other group members (Klemm *et al.*, 1998). These discussion groups provide an ideal opportunity to study potential differences across cancer types in natural written expression about distress-inducing events, including writing about emotions, use of explanatory constructs, and sharing information (Davison *et al.*, 2000; Stone and Pennebaker, 2002).

In the present study, we sought to quantitatively replicate an earlier qualitative study of these online groups (Klemm *et al.*, 1999) by quantifying linguistic patterns of emotional expression, cognitive processing, and cancer-related information exchange in online cancer discussion groups for breast and prostate cancer. In the Klemm *et al.* (1999) study, four investigators reviewed each line of each message posted to the breast and prostate discussion lists and categorized general themes of discussion into one of nine categories, including personal experience, encouragement/support, and information giving/seeking. In the present study, we employed Linguistic Inquiry and Word Count (LIWC), a quantitative text analysis program that calculates the proportion of all words contained in a written sample that fall into specific word categories (e.g., affect-related words and cognitive mechanisms). We also developed cancer-specific word categories to supplement LIWC with the ability to identify words related to medical treatments and settings that are infrequently used in the English language. Consistent with qualitative findings from these online groups (Klemm *et al.*, 1999), we hypothesized that women with breast cancer would exhibit higher rates of interaction and emotional disclosure and lower rates of information sharing than men with prostate cancer.

METHOD

Participants

After obtaining approval from the UAB Institutional Review Board, writing samples were obtained from two cancer-specific discussion groups

made publicly available on the Internet, the Breast Cancer List and the Prostate Problems Mailing List. Both groups have been described in previous reports and subjected to content analysis (Klemm *et al.*, 1999). Messages posted to each group between May 2, 2001 and June 30, 2001 that were in the public domain, were analyzed to evaluate patterns of word use, and only messages deemed to be sent by a cancer patient were included for analysis. Research assistants reviewed the content of each message and included only messages that described first-hand experience with cancer diagnosis, treatment, or survival. All messages sent by family members, friends, or others participating in the discussion lists who identified themselves as a nonpatient were excluded. Messages that were ambiguous with regard to the diagnostic status of the sender were assumed to be written by a patient and included in subsequent analyses. Messages from 325 individuals in the Breast Cancer List and 162 individuals in the Prostate Problems Mailing List were analyzed. Only the written text of messages was kept for analysis. Unique e-mail addresses associated with messages were replaced with a subject code so that multiple messages posted over time could be tracked. No personal or identifying information was collected.

The Breast Cancer List is a semipublic discussion group established in 1994 that uses listserv technology to create a central forum where e-mail messages from patients, loved ones, friends, and care providers are sent and distributed to all subscribers (Sharf, 1997). The list is automated, and anyone who wishes to receive messages from the discussion group can subscribe. Once subscribed, participants receive dozens of e-mail messages each day from other subscribers to the discussion group. The Prostate Problems Mailing List is also a public discussion group that functions in a similar manner to the Breast Cancer List. Both mailing lists are widely cited on prominent cancer-related information sites as among the largest patient-based discussion groups for breast and prostate cancer. At the time this study was undertaken, these lists were archived on the web and in the public domain, but this is no longer true for the prostate list. Eysenbach and Till (2001) have outlined ethical issues associated with passive analysis of Internet communities, suggesting that informed consent can be waived when messages are deidentified, mailing list participants are not inconvenienced or harmed, and the study has been IRB approved. These ethical considerations were met for this study.

Procedure

Using archives of all messages sent to each group in the specified time period, each message posted to either discussion group was saved in a unique file, messages were stripped of all possible identifiers, and all nonrelevant

characters were deleted. Nonrelevant characters included e-mail headings (e.g. "to," "from," or "subject" fields) and text written by other participants that was reproduced in the message as a function of the "reply-to" command. Each message was then subjected to content analysis using the word libraries described below.

Word Libraries

Use of words related to emotional expression and cognitive processing was measured using Linguistic Inquiry and Word Count (LIWC, Pennebaker *et al.*, 2001). LIWC compares each word contained in a text file with words or word stems classified into 74 categories. The LIWC calculates the percentage of the total number of words that are contained in each of the 74 categories. Previous research on psychological processes associated with improved outcome has focused on LIWC categories of 1) *emotion*, including subcategories of positive emotion and negative emotion, and 2) *cognitive mechanisms*, including subcategories of insight and causation (Pennebaker, 1997). In this study, only the categories of affect-related words and cognitive mechanisms were analyzed. Inter-rater reliability of LIWC categories range from 0.86 to 1.00, and external validity of the emotional expression and cognitive processing word categories is good (Pennebaker *et al.*, 2001).

An advantage of the LIWC program is that it allows researchers to incorporate additional libraries to assess words not captured by the LIWC dictionary. Because the LIWC dictionary contains only 2290 commonly-used words in the English language and does not contain words specific to oncology, we developed a study-specific set of categories to measure the exchange of cancer-related information. Using Practical Extraction and Report Language (PERL), a programming language useful for the manipulation of text files, we screened all words in all postings to the two discussion groups against the LIWC dictionary and retained all words or phrases that were not captured by LIWC analysis. This process resulted in 23,376 unique words or word fragments.

Three trained raters with several years of clinical and research experience with cancer patients then reviewed these words to determine which were likely to have been used in the context of discussing experiences related to cancer. Reviewers were blind to the discussion group from which each retained term was derived, and two reviewers categorized each unique word or word fragment as either unrelated or related to cancer. Reviewers identified 2229 words as being cancer-related. Inter-rater agreement on whether a word was cancer-related was 77.3% (1,724 of the 2229 words). Words for which there was rater disagreement about inclusion in the cancer libraries were then reviewed by two of the authors (JO, DT), who

discussed each word and reached consensus about eventual placement in the libraries.

Words selected for inclusion in the cancer information libraries were then further categorized into one of seven word categories: outcomes of cancer treatment (e.g., alive, conquered, cured), disease status and processes associated with the cancer (e.g., diagnosed, spread, stage), health care facilities and personnel (e.g., check-up, pharmacist, oncologist, nurse), medical tests and procedures (e.g., imaging, biopsy, scan, self-exam), cancer treatment (e.g., anti-emetic, drug, hysterectomy), physical symptoms and side-effects (e.g., impotence, sterile, bald), and descriptions of cancer in the body (e.g., baseball-sized, metastases, carcinoma, DCIS). Three additional categories were developed to include all U.S. Food and Drug Administration (FDA)-approved drugs, FDA-approved oncology medications, and complementary and alternative medicines. Comprehensive listings of FDA-approved drugs and approved oncology medications were obtained from the FDA web site (http://www.fda.gov). Terms specific to complementary and alternative medicines used by persons with cancer were obtained from a comprehensive textbook listing descriptions of known nontraditional cancer treatments (American Cancer Society, 2000). This process yielded 4547 words in the FDA-approved drugs category (e.g., albamycin, estradiol, narcan, Zyban), 262 words in the FDA-approved oncology medications category (e.g., taxotere, neupogen, fluorouracil, raloxifen), and 894 terms in the complementary and alternative medicines category (e.g., aromatherapy, biofeedback, qigong, capsaicin, greasewood). A total of 7670 unique words, abbreviations, and misspellings of cancer-related terms were retained by the authors. A general cancer library was obtained by collapsing each of the categories of cancer-related words into a single cancer information category. Table I provides an example of LIWC and cancer word categories as applied to fictional sample messages from each group.

Data Analysis

To compare LIWC categories between groups, word categories were calculated as mean % of words used across all messages submitted by individual participants. Moreover, *t* tests were used to compare the two groups on average levels of word use collapsed over the observation period; repeated measures analysis of variance was used to perform longitudinal analyses. We also examined the raw number of words used in each category in order to minimize the potential for certain types of communication to be diluted when messages were longer (e.g., a lengthy message containing a section of strong emotional expression followed by a section of nonemotional

Table I. LIWC and Cancer Library Word Identification Demonstrated in Fictional Text Examples Similar to Those Posted in Each of the Internet Discussion Groups

Discussion group	Text example
Breast	*After*[10,37] *my*[1,2,4] diagnosis[73] *and*[44] *lumpectomy*[60,61], *I*[1,2,4] *had*[38] several treatments[76] *of*[10] *the*[9] AC[70] *and*[44] taxotere[70,71] *combination*[65]. *Although*[45] *my*[1,2,4] oncologist[74] prescribed[76] *something*[1,25] *to*[10] *prevent*[20,24] *nausea*[60,61], *the*[9] onc[74] nurse[74] forgot *to*[10] *give*[31,32,39] *me*[1,2,4] *the*[9] anti-nausea[76] meds[76] *during*[37] *the*[9] *first*[11,47,50] *two*[11] treatments[76]. *I*[1,2,4] *became*[20,22,38] *very*[26] sick[77] *after*[10,37] each chemo[76]. *When*[37] *I*[1,2,4] *realized*[20,22] *that*[45] *I*[1,2,4] *was*[38] *supposed*[25,38] *to*[10] *be*[40] *taking*[46] Zofran[70], *I*[1,2,4] *felt*[20,22,27,30,38] much *better*[12,13,47,50], *and*[44] *the*[9] pills[76] *were*[38] *easy*[12,13,15] *to*[10] *take*[39,46] *with*[10,44] *no*[7] side-effects[77]. *I*[1,2,4] actually tolerated *the*[9] chemo[76] *very*[26] well, *but*[20,23,45] radiation[76] *was*[38] *challenging*[12,13,15,47,50] *for*[10] *me*[1,2,4]. *I*[1,2,4] *felt*[20,22,27,30,38] *helpless*[12,16,19] *against*[10] *the*[9] *fatigue*[12,16,19]60,61,64, *I*[1,2,4] *started*[37,38] gaining[77] *weight*[60,63], *and*[44] *I*[1,2,4] *was*[38] *just*[25] *too*[44] *tired*[60,61,64] *to*[10] *exercise*[51,53]. *I*[1,2,4] *became*[20,22,38] *very*[26] *depressed*[12,16,19] *for*[10] *a*[9] *while*[37]. *Thankfully*[12,13,14], *my*[1,2,4] *husband*[31,35] *was*[38] *very*[26] *supportive*[12,13], *and*[44] *my*[1,2,4] *friends*[31,34] rallied *around*[10,41] *me*[1,2,4]. *Without*[7,10,45] *them*[1,6,31,33] (*and*[44] *all*[26] *of*[10] *you*[1,5,31,33]), *I'm*[1,2,4,39] *not*[7] *sure*[12,13,26] *I*[1,2,4] *would*[20,23] *have*[39] survived[72!] *I*[1,2,4] *have*[39] *now*[37] *graduated*[47,48] *from*[10] *poison*[12,16,18] *and*[44] beams *to*[10] tamoxifen[70,71], yoga[69], *and*[44] acupuncture[69].
Prostate	*Prostate*[60,61,62] *cancer*[60,61] *for*[10] *me*[1,2,4] *began*[37,38] *when*[37] *my*[1,2,4] *annual*[37] physical *exam*[47,48] *yielded*[20,24] *a*[9] PSA[75] *of*[10] 85. *I*[1,2,4] elected *not*[7] *to*[10] *have*[39] *a*[9] *prostatectomy*[60,61,62], *thinking*[20,22] *that*[45] *if*[10,20,23] *I*[1,2,4] *did*[38] brachytherapy[76] *that*[45] *it*[1] *would*[20,23] spare *me*[1,2,4] *some*[25] *of*[10] *the*[9] problems *I've*[1,2,4] *heard*[27,29,31,32,38] *about*[10]. *I*[1,2,4] *haven't*[7,39] *had*[38] *any*[25] problems *that*[45] *some*[25] *of*[10] *you*[1,5,31,33] *have*[39] *described*[20,31,32,38] *with*[10,44] leaking[77] *or*[25,45] erectile dysfunction[72]. *Thankfully*[12,13,14], *my*[1,2,4] *bone*[60,61] scan[75] *was*[38] *clean*[51,52,60,65] (*no*[7] metastasis[78]), *and*[44] *my*[1,2,4] urologist[74] *has*[39] put *me*[1,2,4] *on*[10,42] Lupron[70,71] *and*[44] Casodex[70,71]. *The*[9] only thing *I*[1,2,4] *worry*[12,16,17] *about*[10] *is*[39] *the*[9] *occasional*[25,37] hot flash[77]. *The*[9] experts[74] *are*[39] *saying*[27,29,31,32] *that*[45] *I'm*[1,2,4,39] *in*[10,44] remission[72], *and*[44] *my*[1,2,4] chances[73] *are*[39] *good*[12,13] *for*[10] *a*[9] *pretty*[12,13] normal life[72]. *I*[1,2,4] *can't*[7,39] *say*[27,29,31,32] *that*[45] *I*[1,2,4] *regret*[12,16] *any*[25] *of*[10] *the*[9] decisions *that*[45] *I*[1,2,4] *made*[38] - *but*[20,23,45] *they*[1,6,31,33] *were*[38] *difficult*[12,16] decisions *to*[10] *make*[39].

Note. [*Italics* represents words identified by LIWC; Underline represents words identified by supplemental cancer libraries] *LIWC Word Libraries*: 1 = Total Pronouns; 2 = First Person Singular; 3 = First Person Plural; 4 = Total First Person; 5 = Total Second Person; 6 = Total Third Person; 7 = Negations; 8 = Assents; 9 = Articles; 10 = Prepositions; 11 = Numbers; 12 = Affective or Emotional Processes; 13 = Positive Emotions; 14 = Positive Feelings; 15 = Optimism and Energy; 16 = Negative Emotions; 17 = Anxiety or Fear; 18 = Anger; 19 = Sadness or Depression; 20 = Cognitive Processes; 21 = Causation; 22 = Insight; 23 = Discrepancy; 24 = Inhibition; 25 = Tentative; 26 = Certainty; 27 = Sensory and Perceptual Processes; 28 = Seeing; 29 = Hearing; 30 = Feeling; 31 = Social Processes; 32 = Communication; 33 = Other References to People; 34 = Friends; 35 = Family; 36 = Humans; 37 = Time; 38 = Past Tense Verb; 39 = Present Tense Verb; 40 = Future Tense Verb; 41 = Space; 42 = Up; 43 = Down; 44 = Inclusive; 45 = Exclusive; 46 = Motion; 47 = Occupation; 48 = School; 49 = Job or Work; 50 = Achievement; 51 = Leisure Activity; 52 = home; 53 = Sports; 54 = Television and Movies; 55 = Music; 56 = Money and Financial Issues; 57 = Metaphysical Issues; 58 = Religion; 59 = Death and Dying; 60 = Physical States and Function; 61 = Body States and Symptoms; 62 = Sex and Sexuality; 63 = Eating, Drinking, and Dieting; 64 = Sleeping and Dreaming; 65 = Grooming; 66 = Swear Words; 67 = Nonfluencies; 68 = Fillers. *Cancer Word Libraries:* 69 = Complementary and Alternative Medical Treatments; 70 = All FDA-Approved Drugs; 71 = All FDA-Approved Oncology Drugs; 72 = Treatment Outcome; 73 = Tumor Activity and Medical Procedures; 74 = Health-Care Practitioners; 75 = Medical Tests and Diagnostics; 76 = Cancer Treatment; 77 = Physical Symptoms and Side-Effects; 78 = Tumor Characteristics.

expression could result in a lower percentage of total words related to emotion). The pattern of results was identical when absolute word counts and the standard outputs of the LIWC (i.e., percentage of words in categories) were analyzed. Therefore the standard output of LIWC is reported in this article. In analyses demonstrating group differences in use of words contained within the general cancer library, follow-up analyses to test potential group differences in use of words found within each of the eight subcategories (i.e., outcomes of cancer treatment, disease status and processes, health care facilities and personnel, medical tests/procedures, cancer treatment, physical symptoms & side-effects, descriptions of cancer in the body, FDA-approved drugs, and FDA-approved oncology drugs, and complementary and alternative treatments) were employed.

RESULTS

Communication Patterns Within Groups

Using a series of *t* tests, the two cancer discussion groups were compared across several types of communication patterns (see Table II). The Breast Cancer List showed more postings per day and more postings per participant. Among participants who submitted more than one message over the observed sampling period, Breast Cancer List participants exhibited shorter time intervals between postings than did prostate cancer list participants. Additionally, LIWC captured significantly fewer words used by participants in the Prostate Problems Mailing List, whereas the Cancer Information library captured significantly fewer words in messages posted to the Breast Cancer List.

Table II. Characteristics of Communication Within the Discussion Groups

	Breast cancer list $n = 325$		Prostate problems mailing list $n = 162$			
	M	SD	*M*	SD	*F*-value	DF
Number of postings per day	37.9	11.4	6.4	4.1	32.8***	1,59
Number of postings per participant	7.0	8.4	2.4	2.9	46.7***	1,485
Length of posting, in words	180.6	159.3	160.4	156.1	1.75NS	1,484
Interval between postings, in days	6.2	8.4	12.3	11.4	24.2***	1,307
% of words captured by LIWC libraries	77.2	6.1	73.6	7.4	32.9***	1,484
% of words captured by Cancer Information Library	4.8	2.9	8.6	6.0	86.7***	1,484

Note. NS = Not a significant difference.
***$p < 0.0001$.

Group Differences in Use of Emotion-Related Words

In order to evaluate the hypothesis that breast cancer patients would exhibit higher levels of emotional disclosure than prostate cancer patients, LIWC categories of words related to positive emotional states, negative emotional states, and all emotional states were compared between the two groups using *t* tests. Across the 2 months of the study sample, messages from breast cancer patients made significantly higher use of words related to emotion, $t(484) = 2.79, p = 0.006$. Both subcategories comprising the emotion-related category were also significantly higher in writing samples from the breast cancer group, including positive emotion, $t\ (484) = 4.95$, $p < 0.0001$, and negative emotion, $t(484) = 2.79, p = 0.006$ (see Fig. 1). Differences between groups in positive emotions were demonstrated by higher use of words related to positive feelings, $t(484) = 5.17, p < 0.0001$, in the breast cancer group. There were no group differences in use of words related to optimism. Differences between groups in use of negative emotion words were manifested by more frequent expressions of anxiety in the breast cancer group, $t(484) = 2.91$, $p = 0.004$. The two groups did not differ significantly in their use of words related to anger or sadness.

Group Differences in Use of Cognitive Words

Use of words related to cognitive mechanisms was significantly higher among participants in the Breast Cancer List, $t(484) = 2.29$, $p = 0.023$. No

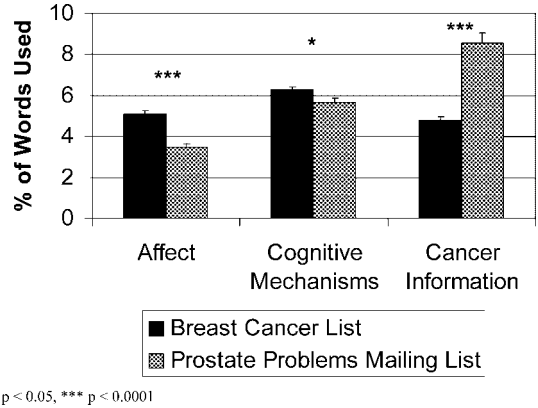

Fig. 1. Mean percentage of words used by category for each discussion group.

group differences were observed in subcategories of words related to causation ($p = 0.058$) or insight ($p = 0.064$). Use of words related to social coping differed significantly between the two groups, with Breast Cancer Discussion List participants exhibiting higher use of words related to social processes, $t(484) = 5.31$, $p < 0.0001$.

Group Differences in Use of Cancer-Related Words

We then sought to compare the two groups of participants on use of words related to physical and medical experiences using the LIWC word category of physical states and functions and the cancer-related word categories developed for this study. No group differences in use of words contained in the LIWC physical states and functions were detected. However, strong group differences were identified in the use of words relating to cancer information, $t(484) = 9.31$, $p < 0.0001$. Participants in the Prostate Problems Mailing List used, on average, nearly twice as many words pertaining to cancer information (see Fig. 1). Significant group differences were also observed for a number of subcategories within the cancer information library. Prostate cancer patients made higher use of words related to disease status, $t(484) = 3.56$, $p = 0.0004$, healthcare facilities and personnel, $t(484) = 3.47$, $p = 0.0006$, medical tests and procedures, $t(484) = 7.48$, $p < 0.0001$, cancer treatment, $t(484) = 3.95$, $p < 0.0001$, description of cancer in the body, $t(484) = 3.49$, $p = 0.0005$, complementary and alternative treatments, $t(484) = 4.85$, $p < 0.0001$, all FDA-approved drugs, $t(484) = 3.25$, $p = 0.001$, and FDA-approved oncology medications, $t(484) = 2.38$, $p = 0.018$. No group differences were observed for words related to outcomes of treatment or physical symptoms and side-effects.

Longitudinal Analysis of Word Use

Given the longitudinal nature of the database obtained from the two discussion groups, we further sought to characterize word use over time among participants who submitted multiple messages ($n = 236$ for the Breast Cancer List; $n = 74$ for the Prostate Problems Mailing List). Because of the small number of repeat messages posted by participants in the prostate group, longitudinal analyses were restricted to two time points. Repeated measures analysis of variance was employed to test changes over time between the first and second messages posted by each participant. There were no significant effects of Time or the Group $\times$ Time interaction for any of the primary variables analyzed: affect-related words, cognitive-processes, or cancer information words.

t tests were used to evaluate the hypothesis that participants who submitted only one message to a group over the course of the observation period ($n = 88$, 27.1%, for the Breast Cancer List, $n = 89$, 54.9%, for the Prostate Problems Mailing List) might be doing so primarily for the purpose of receiving information. Collapsing across groups, participants submitting multiple postings and those submitting only one posting did not differ in the use of affect-related words but did exhibit different patterns in the use of words related to cognitive mechanisms, $t(484) = 2.53$, $p = 0.01$, insight, $t(484) = 2.12$, $p = 0.035$, and cancer information $t(484) = 3.72$, $p = 0.0002$. Participants who submitted only one message to the groups used fewer cognitive-related words (5.6 vs. 6.3%), fewer insight-related words (1.7 vs. 2.0%), and more words related to cancer information (7.1 vs. 5.5%). Because these participants could be using the discussion groups for qualitatively different purposes, we then compared patterns of word use among only participants who had submitted multiple messages to the groups (Breast $n = 236$, Prostate $n = 74$). The pattern of results was unchanged: participants in the Breast Cancer List used a significantly higher proportion of words related to affective and cognitive processes and a significantly lower proportion of words related to cancer information.

DISCUSSION

The results of this study reveal significant differences between breast and prostate cancer patients on several aspects of Internet-based communication. These findings may reflect known differences in the underlying psychosocial needs of these patients (Kiss and Meryn, 2001). While breast and prostate cancer patients are typically first diagnosed within similar age distributions (Ries *et al.*, 1994) and share some psychosocial concerns (e.g., changes in sexuality, side-effects of treatment), there are some notable differences. Importantly, patients diagnosed with breast cancer have more well-defined treatment options than do patients diagnosed with prostate cancer (Carroll *et al.*, 2001; Winer *et al.*, 2001). As a result, prostate cancer patients may use the Internet primarily as a means of obtaining information about treatment options and anticipated outcomes.

Underlying gender differences in emotional expression may also contribute to the observed communication differences between breast and prostate cancer patients. Taylor *et al.* (2000) have proposed that when confronting stressful situations, females may be more likely than males to engage in creating, nurturing, and maintaining social relationships as a means of reducing emotional distress. Several empirical studies lend support to the supposition that men are more reluctant than women to discuss their feelings. In

a study asking men and women to respond to emotionally provocative video segments, males tended to not use emotional words in describing their experience (Sells and Martin, 2001). Further, developmental studies of emotional expression suggest that between early and late adolescence, males tend to exhibit declining levels of emotional expression, whereas females exhibit increasing levels of emotional expression (Brody, 1996; Polce *et al.*, 1998). Additionally, we have shown that college-age males using the Internet to discuss the experience of cancer in a loved one exhibit greater declines in emotional expression over time than do females (Owen *et al.*, 2003). Future studies of gender differences within a mixed-gender diagnostic category (e.g., lung cancer) could clarify the contributing effects of psychosocial concerns and gender.

Our findings also suggest that Internet-based discussion groups differ in several key respects from expressive writing interventions and existing adjuvant psychological therapies for cancer patients. Writing samples from the Internet discussion groups were approximately half the length of writing samples submitted by subjects across 43 emotion-writing studies (327 words per sample; Pennebaker *et al.*, 2001). Messages submitted to the breast cancer group more closely resembled samples from expressive writing interventions: they submitted an average of seven messages (thus approximating the total word count of participants in emotional writing conditions, who average three writing samples), and they exhibited high levels of emotion-based writing (5.1% of total words, compared to 5.3% in expressive writing groups; Pennebaker *et al.*, 2001). Neither breast nor prostate cancer patients, however, exhibited an increase over time in the amount of cognitive processing expressed in their postings. These results suggest that writing samples from Internet cancer discussion groups do not show patterns of word use that have been previously shown to result in improved physical and emotional outcomes. Finally, while LIWC captured an impressive percentage of the total number of words used by the online participants, it did not capture many of the medical and cancer-specific words. Use of cancer-specific libraries significantly improved word detection in these samples.

The findings of the present study suggest that persons diagnosed with breast and prostate cancer use Internet discussion groups differently. While both use these discussion groups to seek and provide emotional and informational support, the relative proportion of written effort allotted to these types of support differs between the two groups. These results may be useful for enhancing recruitment strategies and service delivery for adjuvant psychological treatments, offered online or face-to-face. Prostate cancer patients may be more willing to participate in psychological treatment if the treatment is presented as one that provides information and supports decision-making processes, whereas treatments that offer support systems

and private outlets for emotion may be of more interest to patients with breast cancer. By framing the way treatments are introduced to patients, the overall impact of psychosocial care could be increased (Owen *et al.*, 2004; Owen *et al.*, in press-b). Presumably, the goals of psychological treatment will be similar for both groups of patients—to help patients obtain the informational, emotional, cognitive, and social resources necessary to successfully cope with the challenges and changes associated with cancer. Further, the Smyth (1998) findings that men exhibit greater effect sizes from structured emotional writing are consistent with our results. If increased emotional expression and cognitive processing, as measured by LIWC, accurately reflect psychological changes that can improve patient outcomes, then prostate cancer patients who turn to Internet discussion groups may benefit from preparation, facilitation, or prompting in order to increase rates of emotional expression and cognitive processing of the cancer experience. Group facilitation could also bolster cognitive processing of the cancer experience for breast cancer patients over time. In a recent controlled study of gender differences in online communication, groups of young adult males exhibited increasing levels of cognitive processing over time when communication was facilitated by questions and prompts (Owen *et al.*, 2003).

Given the quasiexperimental design of the present study, there are several notable limitations of the reported findings. Most importantly, basic demographic characteristics of participants were not obtained. Thus group differences in communication may reflect underlying differences in age, education level, socioeconomic status, ethnicity, or familiarity with the Internet. Recent results from a clinical trial of an Internet-based support group for women with early stage breast cancer suggest that age, socioeconomic status, time since diagnosis, treatment status, and disease status are unrelated to message length, use of affect-related words, or use of words expressing cognitive mechanisms (Owen *et al.*, 2003). However, greater educational background is associated with longer messages and higher use of words related to affect. We were also unable to control for group size and frequency of posting. Also, observational studies of Internet-based communication do not allow researchers to determine basic characteristics of participants—including gender of the participant. Participants in this study were identified only by the group in which they were submitting messages, and gender was inferred from their cancer type and by the nature of their remarks. Further, it is likely that the use of affective, cognitive, and information-related words vary with time since diagnosis and stage of treatment or follow-up. Controlled studies of Internet-based communication, in which key patient characteristics such as stage of disease and date of diagnosis can be assessed, will be necessary to address use of these categories of words over time.

Qualitative analysis of communication in Internet discussion groups reveals much about how these forums are currently used by patients and how online communication may differ from written disclosure or transactional communication in the context of face-to-face therapeutic interactions. Importantly, when used with medical populations, such as the cancer patients observed in this study, the capacity of LIWC to classify words used by participants is reduced from an average of 83.3% (Pennebaker *et al.*, 2001) to 73–77% classified. Because patients use written communication to seek and provide information about medical treatments and processes that are not detected by LIWC, it is important to develop population-specific libraries that can capture and assess these types of communication. In so doing, we were able to detect approximately 82% of all words used by participants in both discussion groups. The development of condition- or disease-specific libraries may greatly increase the applicability of LIWC for some populations and allow researchers to employ variables related to language use in more sophisticated ways to understand the coping efforts and psychological needs of persons confronting life-threatening diseases.

ACKNOWLEDGMENTS

The present results are part of a larger ongoing study of the Internet as a source of support and information for women with breast cancer, supported by grant No. DAMD17-00-1-0121 from the Department of the Army Breast Cancer Research Program.

REFERENCES

American Cancer Society (2000). *Complementary and alternative cancer methods*, American Cancer Society, Atlanta, Georgia.

Brody, L. R. (1996). Gender, emotional expression, and parent–child boundaries. In *Emotion: Interdisciplinary perspectives.* Kavanaugh, R. D., and Zimmerberg, B. (Eds.), Erlbaum, Hillsdale, NJ.

Campbell, J. L. (1996). Traditional men in therapy: Obstacles and recommendations. *J. Psychol. Pract.* 2(3): 40–45.

Carroll, P. R., Lee, K. L., Fuks, Z. Y., and Kantoff, P. W. (2001). Cancer of the prostate. In Devita, V. T., Hellman, S., and Rosenberg, S. A. (Eds.), *Cancer: Principles and practice of oncology*, Lippincott Williams & Wilkins, Philadelphia, pp. 1418–1479.

Classen, C., Koopman, C., Angell, K., and Spiegel, D. (1996). Coping styles associated with psychological adjustment to advanced breast cancer. *Health Psychol.* 15(6): 434–437.

Davison, K. P., Pennebaker, J. W., and Dickerson, S. S. (2000). Who talks? The social psychology of illness support groups. *Am. Psychol.* 55(2): 205–217.

Epstein, J., and Klinkenberg, W. D. (2001). From Eliza to Internet: A brief history of computerized assessment. *Comput. Hum. Behav.* 17(3): 295–314.

Esterling, B. A., L'Abate, L., Murray, E. J., and Pennebaker, J. W. (1999). Empirical foundations for writing in prevention and psychotherapy: Mental and physical health outcomes. *Clin. Psychol. Rev.* 19(1): 79–96.

Eysenbach, G., and Till, J. E. (2001). Ethical issues in qualitative research on Internet communities. *BMJ* 323(7321): 1103–1105.
Fogel, J., Albert, S. M., Schanbel, F., Ditkoff, B. A., and Neuget, A. I. (2002). Internet use and social support in women with breast cancer. *Health Psychol.* 21(4): 398–404.
Kiss, A., and Meryn, S. (2001). Effect of sex and gender on psychosocial aspects of prostate and breast cancer. *BMJ* 323(7320): 1055–1058.
Klemm, P., Hurst, M., Dearholt, S. L., and Trone, S. R. (1999). Gender differences on Internet cancer support groups. *Comput. Nurs.* 17(2): 65–72.
Klemm, P., Reppert, K., and Visich, L. (1998). A nontraditional cancer support group: The Internet. *Comput. Nurs.* 16(1): 31–36.
Meric, F., Bernstam, E. V., Mirza, N. Q., Hunt, K. K., Ames, F. C., Ross, M. I., Kuerer, H. M., Pollock, R. E., Musen, M. A., and Singletary, S. E. (2002). Breast cancer on the World Wide Web: Cross-sectional survey of quality of information and popularity of websites. *BMJ* 324(7337): 577–581.
Owen, J. E., Klapow, J. C., Roth, D. L., Nabell, L., and Tucker, D. C. (2004). Improving the effectiveness of adjuvant psychological treatment for women with breast cancer: The feasibility of providing online support. *Psychooncology* 13(4): 281–292.
Owen, J. E., Klapow, J. C., Roth, D. L., Shuster, J. L., Bellis, J., Meredith, R., and Tucker, D. C. (in press, b). Randomized pilot of a self-guided Internet coping group for women with early-stage breast cancer. University of Alabama at Birmingham, Birmingham, AL, unpublished manuscript.
Owen, J. E., Yarbrough, E. J., Vaga, A., and Tucker, D. C. (2003). Investigation of the effects of gender and preparation on quality of communication in Internet support groups. *Comput. Hum. Behav.* 19(3): 259–275.
Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychol. Sci.* 8(3): 162–166.
Pennebaker, J. W. (2000). Telling stories: The health benefits of narrative. *Lit. Med.* 19(1): 3–18.
Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001*, Erlbaum, Mahwah, NJ.
Polce, L. M., Myers, B. J., Kilmartin, C. T., Forssmann-Falk, R., and Kliewer, W. (1998). Gender and age patterns in emotional expression, body image and self-esteem: A qualitative analysis. *Sex Roles* 38(11–12): 1025–1048.
Ries, L. A. G., Miller, B. A., Hankey, B. F., Kosary, C. L., Harras, A., and Edwards, B. K. (1994). *SEER cancer statistics review, 1973–1991: Tables and graphs*, NIH Pub. No. 94-2789, National Cancer Institute, Bethesda, MD.
Sells, D. J., and Martin, R. B. (2001). Gender and modality differences in experiencing and emotional expression. *Can. J. Couns.* 35(2): 176–188.
Sharf, B. F. (1997). Communicating breast cancer on-line: Support and empowerment on the Internet. *Women Health* 26(1): 65–83.
Smyth, J. M. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *J. Consult. Clin. Psychol.* 66(1): 174–184.
Stanton, A. L., Danoff-Burg, S., Cameron, C. L., Bishop, M., Collins, C. A., Kirk, S. B., Sworowski, L. A., and Twillman, R. (2000). Emotionally expressive coping predicts psychological and physical adjustment to breast cancer. *J. Consult. Clin. Psychol.* 68(5): 875–882.
Stone, L. D., and Pennebaker, J. W. (2002). Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana. *Basic Appl. Soc. Psychol.* 24(3): 173–183.
Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurun, R. A. R., and Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychol. Rev.* 107(3): 411–429.
Winer, E. P., Morrow, M., Kent, O. C., and Harris, J. R. (2001). Malignant tumors of the breast. Cancer of the prostate. In Devita, V. T., Hellman, S., and Rosenberg, S. A. (Eds.), *Cancer: Principles and practice of oncology*, Lippincott Williams & Wilkins, Philadelphia, pp. 1654–1717.