Breast cancer stories on the internet : improving search facilities to help patients find stories of similar others
Overberg, R.I.

Citation
Overberg, R. I. (2013, September 10). *Breast cancer stories on the internet : improving search facilities to help patients find stories of similar others*. Retrieved from https://hdl.handle.net/1887/21706

Cover Page

The handle http://hdl.handle.net/1887/21706 holds various files of this Leiden University dissertation.

Author: Overberg, Regina Ingrid
Title: Breast cancer stories on the internet : improving search facilities to help patients find stories of similar others
Issue Date: 2013-09-10

SUMMARY

The primary aim of this thesis is to gain insight into which search facilities for spontaneously published stories facilitate breast cancer patients in finding stories by other patients in a similar situation. Spontaneously published stories are defined as first-person texts written by breast cancer patients in which these patients describe their experiences with the illness trajectory, their feelings about having breast cancer, and how they cope with the disease, and which are published at a given moment on the Internet as a 'completed' story (i.e. without later additions and not as part of an interactive forum). The research was conducted from a clinical informatics perspective, using knowledge from the fields of information retrieval, human-computer interaction and psychology.

The rationale behind the study is that spontaneously published breast cancer stories on the Internet are a unique source of support and information for other breast cancer patients, since these stories are mostly complete stories, the writers are not selected, the stories are not edited, and the number of these stories is large and continuously increasing. According to the narrative approach, social comparison theory, and social cognitive theory, reading stories about similar others may have the most positive impact. Breast cancer was chosen because of its relatively high prevalence and its impact on physical, psychological, and social levels.

The research followed a user-centred design: users of search facilities (i.e., patients who want to read stories written by others) were involved in every phase of the research so that the search facilities could be tailored to these users' needs and preferences instead of forcing users to change their search behaviour to conform to search facilities that were developed without their involvement. Concretely, we examined which breast cancer patients spontaneously publish their story online and which themes they write about in their stories (*Chapter 3*), and how these stories are presented and disclosed on websites (*Chapter 2* and its *addendum* containing an update conducted in 2013). We then examined which topics and writer characteristics patients prefer to search for (*Chapter 4*) and translated these preferences into search facilities. Finally, we examined which search facilities show the most positive effects on patient satisfaction and search success, and how patients actually search for stories written by others (*Chapters 5-7*).

Spontaneously published breast cancer stories on the Internet

The study in the addendum to *Chapter 2* shows that on six out of the sixteen websites the stories had been edited by an editor, and that only for three websites could it be confirmed that part of the stories had actually been written or recounted by real patients. Whether stories are reviewed by an editor or written/told by real patients might affect the reliability and authenticity of the stories. However, story reliability and authenticity are difficult to define and the question is how these aspects affect readers of stories. Furthermore, the addendum to Chapter 2 shows that only two websites offer a search facility based on writer characteristics (for example: age, treatment undergone, etc.) and none of the examined websites offer a search facility for story topics based on content keywords.

The results of the study in *Chapter 3* point to a spontaneous selection of both breast cancer patients who publish their story on the Internet and the themes about which they write: the patients and the themes do not exactly reflect the larger population of breast cancer patients and the prevalence of problems they are facing. This result raises questions about the representativeness of a set of spontaneously published stories and the importance of representativeness for the reader.

When the results of *Chapters 2 and 3* are considered together, the question that arises is how a reader can be best helped in easily finding stories by other patients in a similar situation and which may therefore offer the most support and information. To answer this question, *Chapters 4-7* examine several aspects of search facilities for breast cancer stories.

Search facilities for stories

The interviews with breast cancer patients in *Chapter 4* showed that these patients preferred to read about the following themes in other patients' stories: coping with emotions, the impact of cancer on everyday life, and physical discomforts. Patients wanted to be able to select stories according to disease and demographic writer characteristics. The three most preferred writer characteristics were treatment undergone, age and presence of metastases. Most patients preferred part of a story about a specific theme to be presented in text format. Some of them wanted to be able to click on the complete story or on sound and/or video clips.

Chapter 5 describes an online randomised controlled experiment in which we implemented two of the user requirements found in Chapter 4 were implemented: a search facility for story topics and a search facility for writer characteristics. The experiment followed a 2x2 factorial design. The two factors were: searching for story topics 'yes/no' and searching for writer characteristics 'yes/no'. This resulted in four groups: 1) control group (no search facility); 2) search facility for story topics; 3) search facility for writer characteristics; and 4) combination group (a search facility for story topics as well as a search facility for writer characteristics). The results showed that, compared to those

patients who were not able to search for story topics, breast cancer patients who were able to search for topics in the stories scored more positively in terms of satisfaction with the search process, satisfaction with the stories retrieved and self-reported impact of the stories on coping with breast cancer. Searching for writer characteristics resulted in fewer positive effects and only on two of the three outcome measures (satisfaction with the search process and satisfaction with the stories retrieved). Moreover, combining the two search facilities barely amplified the effect. Based on the results of Chapter 5, we can conclude that a search facility for story topics is the most favourable way to disclose breast cancer stories.

The analysis of the log data from this experiment (*Chapter 6*) yielded a contradictory result: the search facility for writer characteristics was used more intensively than the search facility for story topics. At least two thirds of the participants searched for each of the writer characteristics at least once in their search session. When participants searched for writer characteristics, they tended to search for characteristics that matched their own (social comparison with similar others). For example, in all the searches for 'breast conserving therapy' and in 96% of the searches for 'mastectomy', the participants themselves had undergone these treatments, and in 79% of the searches for 'age at diagnosis' participants searched for their own age at diagnosis. Per search, a mean number of 7 writer characteristics were searched for (out of a maximum number of 13 writer characteristics). With respect to topics, participants searched for an average of 4 topics per search (out of a maximum number of 36 topics). However, for most of the searches carried out, one or two topics were searched for.

Thus, it would seem that both searching for topics and searching for writer characteristics are valuable to the reader (*Chapters 5 and 6*). One relatively time- and cost-effective search facility that would enable simultaneous searches for both writer characteristics mentioned in stories and topics covered by the stories is Latent Semantic Indexing (LSI). This mathematical technique positions a set of stories in a multidimensional space. The position of a story in the space is determined by how often certain words appear in the story and in how many stories out of a set of stories these words appear. The user types a search query in natural language that is also positioned in the multidimensional space: the distance between the search query and a story provides information about the degree of similarity between them. *Chapter 7* reports on an exploratory study in which 24 breast cancer patients searched for stories using the LSI search facility. These patients reported being quite satisfied with the features offered by the LSI facility (typing natural language queries, refining queries after accessing stories). However, the facility was associated with certain difficulties: fine-tuning the multidimensional space took more time and effort than anticipated, and one participant indicated that she saw no relationship between her search query and the stories retrieved by LSI. In conclusion, more research is needed before LSI can be used as a search facility for finding relevant stories.

Conclusion and discussion

The Internet has a role in sharing illness stories: there are patients who spontaneously publish their story on the Internet and there are other patients who read these stories. Such spontaneously published stories on the Internet provide both opportunities and risks.

Opportunities lie in the possibility that stories may serve functional purpose, for both writers and readers, such as offering emotional support or models by which to solve problems. As a result, both writing and reading stories on the Internet should not be regulated: patients who feel the need should be able to publish their story or read stories written by others. The opportunities that online stories offer can be increased by implementing relevant search facilities to search for stories. For example, participants indicated that stories that were retrieved with a 'story topics' search facility had a positive impact on coping with the disease. Stories that are insufficiently disclosed, for example due to a lack of search facilities, will have no impact on a reader.

Possible risks regarding spontaneously published stories on the Internet relate to the representativeness, authenticity and reliability of stories. These concepts are difficult to define with respect to spontaneously published illness stories. However, they are important since they may affect readers. Therefore, readers should be empowered in evaluating spontaneously published stories. This could be done by assigning quality labels to websites containing patient stories or by offering a checklist with criteria that patients can use to assess the quality of such websites.

Besides the 'completed' stories examined in this thesis, there are many new developments on the Internet through which patients can share their illness experiences, such as Twitter and Facebook. Since there will always be patients who want to publish their whole story at once, without interaction with others and without their story being moderated, our expectation is that 'completed' stories will coexist alongside these new developments. However, future research could focus on how features of these new forms (for example, degree of interactivity) affect writers and readers.

It is important that the parties involved (such as patients, organisations that provide websites with stories, health care professionals and patient organisations) discuss the opportunities and risks of spontaneously published stories with each other and that they discuss each others' roles and responsibilities in this context. An effective discussion may help to increase the opportunities and reduce the risks of online patient stories and experiences.