Breast cancer stories on the internet : improving search facilities to help patients find stories of similar others

Overberg, R.I.

APPENDICES

Appendices Chapter 5

Please view Multimedia Appendix 1 and Multimedia Appendix 2 on the Internet. You can find these at:
http://www.jmir.org/article/downloadSuppFile/1215/1405
http://www.jmir.org/article/downloadSuppFile/1215/885

Multimedia Appendix 3 (http://www.jmir.org/article/downloadSuppFile/1215/1406):
An overview of the constructs and items belonging to the three main outcome measures.

Constructs: 1. 'Satisfaction with the search process'
1.a. Opinion about the search facility (Cronbach's alpha = .88):
Five 5-point semantic differentials:
- poor ... good
- difficult ... easy
- not at all helpful ... very helpful
- unpleasant ... pleasant
- not convenient ... convenient

1.b. Opinion about the number of search options (Cronbach's alpha = NA):
Response categories: 'too few', 'quite few', 'not few, not many', 'quite many', 'too many'
- What is your opinion on the number of search options?

(For analysis, this item was recoded into 3 points 1=negative opinion ('too few' or 'too many'), 2=slightly negative opinion ('quite few' or 'quite many'), and 3= positive opinion ('not few, not many')

1.c. The extent to which the search options enable finding the information one was looking for (Cronbach's alpha = .75):
Response categories: 'disagree', 'disagree a bit', 'neither disagree nor agree', 'agree a bit', 'agree'.
The ways in which I could search...
- ...did *not* enable me to find what I wanted to know.
- ...did enable me to find information that is important to me.
- ...did *not* enable me to find stories that were helpful to me.
- ...did enable me to find what I was searching for.

1.d. Recommendation to others' and future own use (Cronbach's alpha = .82):
Response categories: 'yes', 'probably yes', 'maybe yes, maybe not', 'probably not', 'no'.
- Would you recommend others to search for stories in this way?
- Would you yourself search more often for stories in this way?

1.e. Overall satisfaction with the search facility (Cronbach's alpha = NA):
One 10-point semantic differential:
- very poor ... excellent

Constructs: 2. 'Satisfaction with the stories retrieved'
2.a. Opinion about the stories retrieved (Cronbach's alpha = .71):
Six 5-point semantic differentials:
- poor ... good
- difficult ... easy
- not at all helpful ... very helpful
- not informative ... informative
- frightening ... not frightening
- I already knew everything ... everything was knew to me

2.b. Opinion about the number of stories retrieved (Cronbach's alpha = NA):
Response categories: 'too few', 'quite few', 'not few, not many', 'quite many', 'too many'.
- What is your opinion on the number of stories retrieved?
  (For analysis, this item was recoded into 3 points 1=negative opinion ('too few' or 'too many'), 2=slightly negative opinion ('quite few' or 'quite many'), and 3= positive opinion ('not few, not many')

2.c. Opinion about the list of stories displayed after a search (Cronbach's alpha = .76):
Response categories: 'disagree', 'disagree a bit', 'neither disagree nor agree', 'agree a bit', 'agree'.
- It was clear to me which story from the list I had to click onto to find the information I was looking for.
- The list of retrieved stories displayed after a search, was *not* helpful to me.
- It was difficult for me to decide which story I should read first.
- The way in which the list of retrieved stories was displayed, was clear to me.

2.d. The extent to which the stories retrieved covered one's information need (Cronbach's alpha = .82):
Response categories: 'disagree', 'disagree a bit', 'neither disagree nor agree', 'agree a bit', 'agree'.
- I have found the information I was looking for.
- I am *not* satisfied with the information I have found.
- The information satisfied my expectations.
- I did *not* find the information I wanted to know.

2.e. Recommendation to others' and future own reading (Cronbach's alpha = .77):
Response categories: 'yes', 'probably yes', 'maybe yes, maybe not', 'probably not', 'no'.
- Would you recommend others to read these stories?
- Would you yourself read more of these kinds of stories?

2.f. Overall satisfaction with the stories retrieved (Cronbach's alpha = NA):
One 10-point semantic differential:
- very poor ... excellent

Constructs: 3. 'The stories' impact on coping with breast cancer'

3.a. The stories' impact on coping with breast cancer (Cronbach's alpha = .85):
Response categories: 'disagree', 'disagree a bit', 'neither disagree nor agree', 'agree a bit', 'agree'.

Having read the stories:
- ...I have learnt things.
- ...I view things differently.
- ...I am able to understand my feelings better.
- ...I see that certain emotions accompany learning to live with breast cancer.
- ...I know what to do.
- ...I see that others have experienced the same things.

Appendices Chapter 7

Appendix I
Composition of the questions of which the results are presented in this article. The questions come from a questionnaire about search satisfaction that was used in a previous study [1].

- Opinion about the search facility (Cronbach's alpha = .88).
  Five 5-point semantic differentials:
  - poor ... good
  - difficult ... easy
  - not at all helpful ... very helpful
  - unpleasant ... pleasant
  - not convenient ... convenient

- Opinion about the list of stories displayed after a search (Cronbach's alpha = .76).
  Response categories: 'disagree', 'disagree a bit', 'neither disagree nor agree', 'agree a bit', 'agree'.
  - It was clear to me which story from the list I had to click onto to find the information I was looking for.
  - The list of retrieved stories displayed after a search, was not helpful to me.
  - It was difficult for me to decide which story I should read first.
  - The way in which the list of retrieved stories was displayed, was clear to me.

- The extent to which the search options enable finding the information one was looking for (Cronbach's alpha = .75). Response categories: 'disagree', 'disagree a bit', 'neither disagree nor agree', 'agree a bit', 'agree'. The ways in which I could search...
  - o ...did not enable me to find what I wanted to know.
  - o ...did enable me to find information that is important to me.
  - o ...did not enable me to find stories that were helpful to me.
  - o ...did enable me to find what I was searching for.

- Recommendation to others' and future own use (Cronbach's alpha = .82). Response categories: 'yes', 'probably yes', 'maybe yes, maybe not', 'probably not', 'no'.
  - o Would you recommend others to search for stories in this way?
  - o Would you yourself search more often for stories in this way?

- Overall satisfaction with the search facility (Cronbach's alpha = NA). One 10-point semantic differential:
  - o very poor ... excellent

Appendix II

Latent Semantic Indexing [2] was performed on the set of 171 Amazones stories. The ORACLE Dutch stop words list was used to remove stop words [3]. An exception was made for the Dutch word 'haar' meaning 'her', because this word also means 'hair' which is an important word in the context of cancer. The stemming algorithm "Dutch Porter for the UPLIFT project" was used to stem the words [4]. Word stems shorter than three characters were removed. Diacritics were reduced to non-diacritics. TF/IDF was calculated on the basis of the 6241 terms that remained in the 171 stories. The weighted word-by-story matrix consisted of 34343 non-zero values (cutoff=0.0002) and was thus filled for 3% (=34343 /(171 * 6241)). With Singular Value Decomposition (SVD) the number of dimensions was reduced to 12 (eigenvalues greater than 0.32). A quadratic vector normalisation and cosine similarity between stories were used.

References

1. Overberg R, Otten W, de MA, Toussaint P, Westenbrink J, Zwetsloot-Schonk B. How breast cancer patients want to search for and retrieve information from stories of other patients on the internet: an online randomized controlled experiment. J Med Internet Res 2010;12(1):e7.
2. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R. Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science 1990;41(6):391-407.
3. ORACLE Dutch stop words list. ORACLE Dutch stop words list 2011.
4. Kraaij W, Pohlmann RC. Dutch Porter for the UPLIFT project. Research Institute for Language and Speech, Utrecht University; 1994.

CURRICULUM VITAE

Regina Overberg werd geboren op 24 augustus 1978 te Stadskanaal. In 1997 behaalde zij het eindexamen gymnasium B aan het Alexander Hegius Lyceum te Deventer. In datzelfde jaar begon zij haar studie Voeding en Gezondheid aan de Wageningen Universiteit. Deze studie heeft zij afgerond met twee afstudeeronderzoeken: één op het gebied van Humane epidemiologie bij de vakgroep Humane Voeding en Epidemiologie van de Wageningen Universiteit, en één op het gebied van Leefwijze en volksgezondheid bij het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) te Bilthoven. Na haar afstuderen in 2002 heeft Regina een jaar bij het Centrum voor Volksgezondheid Toekomstverkenningen (cVTV) van het RIVM gewerkt. Het promotieonderzoek dat in dit proefschrift beschreven is heeft zij in de periode 2003-2008 bij de groep Klinische Informatiekunde van het Leids Universitair Medisch Centrum (LUMC) uitgevoerd. Sinds 2008 voert zij epidemiologisch onderzoek uit binnen het team Advies & Onderzoek Publieke Gezondheid van GGD Kennemerland te Hoofddorp. Meer informatie is te vinden op: http://nl.linkedin.com/in/reginaoverberg

Regina Overberg was born on the 24th of August 1978 in Stadskanaal, the Netherlands. In 1997 she completed her secondary education at Alexander Hegius Lyceum in Deventer. In the same year she started to study Human Nutrition and Health at the Wageningen University. She graduated in 2002 after having completed two studies: one in the field of human epidemiology at the Department of Human Nutrition and Epidemiology of the Wageningen University, and one in the field of lifestyle and public health at the National Institute for Public Health and the Environment (RIVM) in Bilthoven. After her graduation, Regina worked a year at the Centre for Public Health Forecasting (cVTV) of the RIVM. She performed her PhD research that is described in this dissertation in the period 2003-2008 at the Clinical Informatics group of the Leiden University Medical Centre (LUMC). From 2008, she carries out epidemiological research within the team Advice & Research Public Health at the Municipal Health Service Kennemerland in Hoofddorp. More information is available at: http://nl.linkedin.com/in/reginaoverberg

LIST OF PUBLICATIONS AND PRESENTATIONS*

(in chronological order)

Overberg R, Toussaint P, Zwetsloot-Schonk B. Illness stories on the Internet: Features of websites disclosing breast cancer patients' illness stories in the Dutch language. *Patient Education and Counseling* 2006;61:435-442.

Hoenkamp E, **Overberg R**. Computing latent taxonomies from patients' spontaneous self-disclosure to form compatible support groups. *Studies in Health Technology and Informatics* 2006;124:969-974.

Wolf L, **Overberg R**, Toussaint P, Hoenkamp E, Reckman H. Design of the Narrator system – processing, storing and retrieving medical narrative data. *Journal of Integrated Design & Process Science* 2006;10(4):13-33.

Overberg RI, Campagne AE, Zwetsloot-Schonk JHM. Empowering patients – Use and appreciation of a web application that enables patients to keep their own medical data and to share illness experiences with fellow patients. Poster presentation at the *International Conference on Communication in Healthcare 2006 (EACH)*, 5-8 September 2006, Basle, Switzerland.

Overberg RI, Alpay LL, Verhoef J, Zwetsloot-Schonk JHM. Illness stories on the Internet: What do breast cancer patients want at the end of treatment? *Psycho-Oncology* 2007;16:937-944.

* This is a selection related to the content of this dissertation

Alpay LL, **Overberg RI**, Zwetsloot-Schonk JHM. Empowering citizens in assessing health related websites: a driving factor for healthcare governance. *International Journal of Healthcare Technology and Management* 2007;8:141-160.

Overberg R. Search facilities driven by patients' preferences for breast cancer stories on the Internet. Oral presentation at the *International Conference on Communication in Healthcare 2008* (EACH), 2-5 September 2008, Oslo, Norway. As part of the symposium Web-Based Interventions and Information Technology.

Overberg R, Otten W, De Man A, Toussaint P, Westenbrink J, Zwetsloot-Schonk B. How breast cancer patients want to search for and retrieve information from stories of other patients on the Internet: an online randomized controlled experiment. *Journal of Medical Internet Research* 2010;12(1):e7.

Overberg R, De Man A, Wolterbeek R, Otten W, Zwetsloot-Schonk B. Spontaneously published illness stories on a website for young women with breast cancer: do writers and themes reflect the wider population? *Informatics for Health and Social Care* 2013; 38(1): 54-66.

Overberg R, Otten W, De Man A, Zwetsloot-Schonk B. Searching for breast cancer stories online: What topics and/or writer characteristics do patients search for? *Submitted*

DANKWOORD

Allereerst wil ik graag alle patiënten bedanken die aan de onderzoeken in dit proefschrift hebben meegedaan voor hun interesse, tijd en energie. Zonder hen had ik de onderzoeken niet kunnen uitvoeren. De deelnemers die ik heb geïnterviewd wil ik in het bijzonder bedanken voor de persoonlijke verhalen die ze met mij hebben gedeeld. Daarnaast wil ik de Borstkankervereniging Nederland en Stichting Amazones bedanken voor hun betrokkenheid bij mijn onderzoek. Veel dank gaat uit naar Judith Westenbrink en Diana Gosens die toestemming gaven om de website van Stichting Amazones te gebruiken voor het online experiment naar zoekmogelijkheden naar verhalen. Verder wil ik Tineke van Dijk bedanken voor het feit dat ik haar gouache voor de voorzijde van mijn proefschift mocht gebruiken.

Voor de inspirerende gesprekken en overleggen in het kader van mijn onderzoek wil ik graag mijn promotor prof. dr. Bertie Schonk en mijn co-promotor dr. Wilma Otten bedanken. Ook alle mede-auteurs en oud-collega's van Klinische Informatiekunde wil ik bedanken voor hun ideeën en suggesties: Amy Ahluwalia, Laurence Alpay, Nicole Ezendam, Watze Hoekstra, Eduard Hoenkamp, Lieke Lim, Herman Lodder, Pieter Toussaint, Jan Vanderschoot, John Verhoef, Leo Wolf en Ron Wolterbeek. In het bijzonder wil ik Andries de Man bedanken voor het bouwen van de website voor het online experiment en de LSI-zoekfaciliteit en voor het in goede banen leiden van de enorme databestanden die deze onderzoeken opleverden. Ik heb daarnaast veel gehad aan de frisse blik van Henriët Werij die ik begeleid heb bij haar afstudeeronderzoek en van Kristel Schaap die de verhalen heeft gecodeerd. Verder wil ik Henk Herman Nap en Hilke Reckman bedanken voor de samenwerking binnen het Narrator-project van NWO.

Ook buiten het LUMC heb ik met veel mensen ideeën kunnen uitwisselen over de inhoud en voortgang van mijn onderzoek, die ik daar graag voor wil bedanken: de huidige collega's van GGD Kennemerland, in het bijzonder Jolanda ten Brinke, Resi Cluitmans, Saskia Mérelle, Claire Morssink en Arie Oosterlee, en, vrienden en oud-studiegenoten, in het bijzonder, Marije Brouwer, Daphne Cupido, Carolien van Loo-Bouwman, Anouk Pijpe en Bart van Praag. Speciale dank gaat in dit kader uit naar Joost Klaver.

Tot slot wil ik mijn familie bedanken voor hun vertrouwen, hulp en steun: mama en papa, Gisela en Gerhardus, zus, zwager en lief klein nichtje, Ingrid, Frank en Charli. Bedankt dat jullie er altijd voor me zijn.