



UvA-DARE (Digital Academic Repository)

Different platforms for different patients' needs: Automatic content analysis of different online health information platforms

Sanders, R.; Linn, A.J.; Araujo, T.B.; Vliegenthart, R.; Eenbergen, M.C. van; van Weert, J.C.M.

Published in:
International Journal of Human-Computer Studies

DOI:
[10.1016/j.ijhcs.2019.102386](https://doi.org/10.1016/j.ijhcs.2019.102386)

[Link to publication](#)

License
Article 25fa Dutch Copyright Act

Citation for published version (APA):
Sanders, R., Linn, A. J., Araujo, T. B., Vliegenthart, R., Eenbergen, M. C. V., & van Weert, J. C. M. (2020). Different platforms for different patients' needs: Automatic content analysis of different online health information platforms. *International Journal of Human-Computer Studies*, 137, [102386].
<https://doi.org/10.1016/j.ijhcs.2019.102386>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<http://dare.uva.nl>)



Different platforms for different patients' needs: Automatic content analysis of different online health information platforms



Remco Sanders^{a,*}, Annemiek J. Linn^a, Theo B. Araujo^a, Rens Vliegthart^a,
Mies C. van Eenbergen^b, Julia C.M. van Weert^a

^a Amsterdam School of Communication Research, University of Amsterdam, the Netherlands

^b Department of Research, Netherlands Comprehensive Cancer Organisation, Utrecht, the Netherlands

ARTICLE INFO

Keywords:

Information seeking behavior
Online health information
Information needs
Cancer
Automatic content analysis
Model of patients' social support needs

ABSTRACT

Prior online health research has mainly focused on the predictors or outcomes of online health information, leaving online health information itself understudied. Therefore, online health information has remained an umbrella term encompassing different platforms (expert- vs. peer-generated). A hybrid method that combines qualitative and computational methods is used to identify different topics discussed on these different platforms, and an initial model of patients' social support needs was developed and applied to data obtained from the hybrid method. Using topic modeling ($N_{posts} = 52,990$), topics on two expert- and two peer-generated platforms were identified. Differences between and within platforms were found. While peer-generated platforms mainly covered interaction on emotional support topics, expert-generated platforms covered informational topics. Within peer-generated platforms, patients used their experiences differently.

1. Introduction

The diagnosis of cancer causes patients and their surroundings a great amount of fear and uncertainty (Leydon et al., 2000; Sanson-Fisher et al., 2000). As a result, cancer patients and their relatives have many questions like “what is cancer?” or “what sort of treatment is available?”. Some of these questions get answered by a medical expert. However, many patients experience unmet cognitive and affective needs during their cancer journey, even after a discussion with a medical expert (hereafter called “social support needs”; Cotten and Gupta, 2004; Eysenbach et al., 2004; Puts et al., 2012; Sillence et al., 2006; Iconomou et al., 2001). Acquiring health information online helps patients fulfill these social support needs (Goldsmith, 2004). Currently, up to 70% of chronically ill patients seek online health information (Abdul-Muhsin et al., 2015; Rupert et al., 2016).

Previous research focused either on the predictors or on the outcomes of using online health information (e.g., Cotten and Gupta, 2004; Puts et al., 2012; Sanders and Linn, 2018). However, we lack information on which content, if any, fulfills which social support needs. This information is important to have because online health information has the potential to “set the agenda” for future behavior (Anderson et al., 2003; McCombs and Shaw, 1972; Xiao et al., 2014). A first step is to disentangle the term *online health information*. This term has been used

as an umbrella term to describe multiple platforms on the internet such as online blogs, medical libraries, question-and-answer pages, and discussion forums. These different platforms are likely to contain different topics and are expected to fulfill different social support needs (Anderson et al., 2003; Prestin et al., 2015). While scholars have advocated for a comparison between these platforms, to the best of our knowledge, no previous studies have investigated if and how content differs on these platforms (Coulson et al., 2007; Klemm et al., 2003). Because online health information is often unreliable or difficult to understand, patients and medical experts should work together to fulfill patients' social support needs (Rains and Young, 2009; Buis and Whitten, 2011; Rutten et al., 2005). Knowledge about the content of different platforms helps in this regard by deciding which platform to visit to fulfil specific support needs.

Using a hybrid method that combines a literature review and a computational social science method, the current study aims to investigate the differences in content between online health information platforms (Shah et al., 2015). A theoretical model is developed to serve as a roadmap to identify emerging topics and gain a better understanding of how online platforms relate to patients' social support needs.

The current study adds to the literature in multiple ways. On a methodological level, this study bridges the gap between computational

* Corresponding author.

E-mail address: r.sanders@uva.nl (R. Sanders).

methods and “classic” social science methods. The current study distinguishes itself by using a theoretical model as a starting point. On a theoretical level, this study integrates the diversity of various social support needs and makes a systematic and holistic analysis of the field of online health information research possible. We will demonstrate the usefulness of using a model of social support needs as a basis for unsupervised machine learning. On a practical level, the results allow us to understand which social support needs patients seemingly fulfill online instead of, or in addition to, fulfilling during appointments with a medical expert.

1.1. Towards a model of social support needs

The use of online health information is increasing, and this trend can be placed within a larger societal trend of people taking more control over all aspects of their lives (Beck et al., 1994; Lupton, 1997, 2013). Online health information-seeking behavior can be defined as, “... active efforts to obtain specific information in response to a relevant event” (Niederdeppe et al., 2007, p.154). In the current study, “the relevant event” is operationalized as being diagnosed with or living with (someone with) cancer. The “active effort” is operationalized as searching online for social support. A model will be developed that acknowledges that patients go online because of different social support needs. In this model, it is assumed that due to these different needs, patients search online to find different forms of social support on a variety of topics. To illustrate, users can fulfill their “abstract” cognitive and affective needs (layer I) by going online to find different forms of social support (layer II) which is provided to them via content covering different topics. The layers of the proposed model go from broad to specific. Previous research shows that patients go through similar phases during their health information-seeking process; normally the patient starts with a broad need (e.g., I want to know more about my disease, i.e., cognitive need) and this need becomes more specific through searching and browsing until that specific need is fulfilled (e.g., I want to know more about medicine X; Wilson et al., 2002). See Table 1 for the proposed model of patients’ social support needs.

1.2. Layer I: patients’ needs

The first layer of the framework consists of two basic underlying patients’ needs: *cognitive* and *affective* (Bensing and Verhaak, 2004; Coulson, 2005). Patients have the need to know and understand (i.e., cognitive needs) as well as the need to feel acknowledged and understood (i.e., affective needs). An example of a cognitive need is the need for information about treatment. An affective need can be the need to receive reassurance. This classification is the first layer of the proposed model.

Table 1
Model of patients’ social support needs.

Layer I: Needs	Layer II: Social support category	Sublayer social support topic	Description	
Cognitive needs	Informational support	Cancer-specific information	Information about (diagnosis of) cancer, physical effects, test, symptoms.	
		Treatment-related information	Information about available treatments, how treatment works and side effects.	
		Prognosis information	Change of cure, survival expectations, expectations for future health.	
		Rehabilitation information	Self care, nutrition, availability medical supplies for at home	
		Surveillance and health information	Maintaining or restoring physical health after treatment and preventive behaviors.	
		Interpersonal /social information	Effect of disease on relations with family, friends, caregiver, employment, and social life in general.	
		Financial / legal information	Costs associated with disease (e.g., treatment, insurance, writing a testament).	
Affective needs	Esteem support	Medical system information	Health care system, interactions with health care providers, evaluation of physicians.	
		Body image / sexuality information	Sexuality and physical appearance (i.e., attractiveness).	
		Tangible assistance	Providing, or offering to provide, information, goods, and services.	
		Coping and and sss spiritual	Communicating respect and confidence in ones abilities to deal with disease (e.g., compliments and validation).	
		Network support	Member interaction	Communicating belonging to a group of persons with similar concerns or experiences
			Emotional support	Communicating love, concern, empathy, expressions of emotions or caring, attentive listening, reassurance

1.3. Layer II: social support categories

The cognitive and affective needs specified in layer I can be fulfilled by seeking for and receiving online social support, which is the second layer of the framework. The Social Support Behavior Code and the Optimal Matching Model (SSBC; Cutrona, 1990; Cutrona and Russell, 1990; Cutrona and Suhr, 1992; House, 1983) provide a useful categorization of how patients’ needs can be fulfilled by providing and receiving social support in an online context. Within social support, different categories can be distinguished, namely, *informational support* (fulfilling patients’ cognitive needs) and *esteem, network, and emotional support* (fulfilling patients’ affective needs). These categories are explained below.

Patients’ cognitive needs in layer I can be fulfilled by providing informational support. Informational support means providing guidance to users on topics such as the disease, symptom interpretation, and illness management (Coulson et al., 2007). Rutten et al. (2005) offer a typology that can be used to further specify the social support categories into social support topics that can be measured using topic modeling. This typology is based on a review clustering the support needs of cancer patients about certain health topics or issues. The typology describes nine cognitive informational support topics: *cancer-specific, treatment-related, prognosis, rehabilitation, surveillance and health information, interpersonal/social information, financial/legal information, medical system information, and body image information*. Based on other offline literature *tangible assistance* was added to informational support (Cutrona, 1990). For descriptions of the categories, see Table 1. These informational support categories and topics form the second and third layers of the model.

Rutten et al. (2005) only emotional support topic is *coping information*. The subcategories of coping information given by Rutten and colleagues (2005) seem to include all social support categories (e.g., emotional reactions, support from other patients, and spiritual support). Although previous studies show a correlation between the various support concepts, we believe there is also a clear theoretical distinction between network and esteem support and to a lesser degree, emotional support. Therefore, we use the definition of Rutten and colleagues (2005) in which the coping information category specifies the three forms of support. The following distinction is made: the subtopic *coping and spiritual support* reflects esteem support, the subtopic *member interaction* reflects network support, and the subtopic *emotional support* reflects emotional support. Esteem support refers to communicating respect and confidence in one’s ability to deal with the disease. Network support entails communicating one’s belonging to a group of people sharing the same concerns or (health) complaints. Emotional support refers to messages expressing empathy, love, and

caring (Coulson et al., 2007). Rutten's typology can be used to further specify the social support categories *esteem*, *network*, and *emotional support* into observable social support topics such as coping, spiritual information, and member interaction. See Table 1 for the complete model and the subtopics descriptions.

1.3.1. Towards testing the model of patients' social support needs

By grouping and combining the previously described layers, the model of patients' social support needs is proposed (see Table 1). This model allows us to measure latent needs and social support categories via observable social support topics. By combining the outcomes of previous studies, covering multiple sources of information, and targeting different (cancer) populations, the model consistently conceptualizes and identifies different types of social support topics within online health information platforms (Buis and Whitten, 2011). This process will contribute to a greater generalizability of research findings in the future. A first step would be to map the content of these different platforms. The results of this first step can then be applied to the proposed model. The following research question is proposed: *What social support topics can be distinguished online on cancer-related platforms, and to what extent do these topics fit within with the model (RQ1)?*

1.3.2. Online information on different platforms

The internet offers the unique capability to integrate various forms of content on different platforms (i.e., patients' blogs, questions and answers, medical libraries, and discussion forums; Anderson et al., 2003). To fulfill patients' social support needs, information is required that "matches" their information needs (Cutrona, 1990). Most research has focused on one platform or its content, mainly social network sites (e.g., Hether et al., 2014) or online support groups (e.g., Klemm et al., 2003), or the research did not differentiate between platforms (e.g., Gutierrez et al., 2014).

One way of categorizing these different platforms is based on the author of the information (Rupert et al., 2016). The literature recognizes peer-generated (e.g., the author is a fellow patient writing on blogs or discussion forums) and expert-generated information (e.g., the author is a medical expert writing on a (medical) library or answer-and-questions pages by medical experts). Because fellow patients are "uniquely positioned to understand one another in ways that others, including friends or family, may not" (Namkoong et al., 2013, p.170), peer-generated content can provide a description of experiences to which a patient can relate. This situation includes the potential to read about personal experiences with certain treatments, learn practical tips in dealing with side effects, or exchange emotional support (Hether et al., 2014; Prestin et al., 2015; Rupert et al., 2016; Wang et al., 2015). In contrast, expert-generated platforms can provide patients with medically validated, factual information.

While content written by peers and experts can both include social support categories and topics of the proposed model, the extent to which these categories and topics can be found on these platforms is expected to differ. For example, peer-generated platforms are believed to provide higher levels of affective social support categories compared to expert-generated platforms (Wang et al., 2015). However, it is unknown how these differences relate to the social support topics discussed. Therefore, the current research will include both peer- and expert-generated platforms to gain a better understanding of the content discussed among these different platforms. This situation leads to the following research question: *How do peer- and expert-generated cancer-related platforms differ from each other in terms of social support topics (RQ2)?*

2. Method

2.1. Data

The current study used data retrieved from *Kanker.nl*. *Kanker.nl* is

one of the largest and best-known Dutch websites for cancer-related information. The website offers both peer- and expert-generated platforms which users can read and follow as well as respond to. Peer-generated platforms on this website include blogs and discussion groups. On blogs, users can share their stories in multiple blogposts over a longer period of time. Other users can follow their blogposts and comment on them. Discussion groups contain threads about a variety of topics on which users can ask questions, share experiences, and respond to each other. Expert-generated platforms consist of the medical library and expert questions. The medical library gives information on various types of cancer and covers the cancer trajectory from tests and diagnosis until rehabilitation. Expert questions are a mix of peer- (i.e., questions) and expert-generated content (i.e., answers).

The website is well used: in February 2016, *Kanker.nl* had approximately 390,000 visits. For example, patient blogs had 19,000 unique visitors and 130,000 page views. Discussion groups had 15,500 unique visitors with 50,000 page views per month. The data used in the current study cover the period from April 2013 until November 2016. Since the website is often used and offers both peer- and expert generated platforms, it is a valuable case for the current study.

Users are required to register and must provide their name and a valid e-mail address. An automated e-mail is then sent requiring individuals to confirm their registration by clicking on the registration link. Participants of all platforms within *Kanker.nl* gave (standard) consent for using their data for research when they registered. Ethical approval for the current study was provided by the ethical committee of the authors' university: 2016-PC-7547.

3. Analysis

3.1. Topic modeling

Unsupervised computer-assisted content analysis was used to analyze the data without restricting the analysis based on previously defined categories in coding schemes. This technique has several benefits. First, it allows for the analysis of large amounts of data. This method can be used on multiple datasets to compare, unsupervised, the occurrence of topics on different platforms (i.e., without any interference from the researcher). Second, the analysis can be easily replicated by other researchers (Wang et al., 2015).

Latent Dirichlet Allocation (LDA) was used as a topic modeling method using the *Gensim* package (Řehůřek and Sojka, 2010; Blei et al., 2003). Topic modeling is a computer algorithm that identifies latent constructs by word co-occurrence, using word distribution across documents (e.g., individual posts and questions; Jacobi et al., 2016). Based on these documents, topics were created and displayed as clusters of words. These clusters contain the most-used words for each topic (see *preprocessing* for more information). For each topic, the latent meaning (topic) can be (subjectively) inferred from these words. To derive reliable and interpretable topics, several steps need to be taken within LDA.

3.2. Preprocessing

The raw data received from any online source, including the sources used in this study, is "noisy". Therefore, work needs to be performed before analysis can begin (Petz et al., 2014). Data were preprocessed as follows: First, all *html* mark-up and punctuation was removed (e.g., ** and *nsbp*). Second, stop words were removed using a custom made Dutch stop word list (e.g., *en* [and], *te* [too] and *eens* [once/ever]). Third, topic modeling was done. This step resulted in a list of top words, whose latent meanings can be regarded as topics. A list of top words per topic is provided, as well as a score that indicates the weighted presence of the top word per topic (topic loading). Based on these top words and Fig. 1, the outcome of the topic modeling was evaluated on how meaningful, consistent, and discriminative the identified topics were. Fourth, words in top words deemed unnecessary



Fig. 1. Visualization of Topic Modeling
 Subtopic names and proportions: 1, Sharing current status, 14.90%; 2, Sharing previous experience, 13.60%; 3, Cancer treatment information, 10.70%; 4, Emotions and cancer, 8.40%; 5, Member interaction, 7.40%; 6, Treatment side effects, 5.00%; 7, Affective support, 4.90%; 8, Sharing status relatives, 4.90%; 9, Cancer tests, 4.40%; 10, Rehabilitation, 4.30%; 11, Impact cancer (relatives), 3.80%; 12, After care, 2.60%; 13, Cancer organizations, 2.10%; 14, Cancer detection, 2.00%; 15, Insurance/financial, 1.60%; 16, Cancer research, 1.50%; 17, Nutrition, 1.30%; 18, Employment, 1.30%; 19, Other (alternative) treatments, 1.00%; 20, Medical trials, 0.90%; 21, Gastrointestinal treatment, 0.90%; 22, Hereditary, 0.80%; 23, Treatment consequences, 0.80%; 24, Information sources, 0.50%; 25, External organizations, 0.50 %.

or uninformative were removed by adjusting the code. For example, based on the first and second cycles, all names were replaced by the word *naam* [name] using a custom made list of popular Dutch names. The same procedure was repeated for all forms of cancer, which were replaced by the word *kankervorm* [type of cancer]. This process was used since an individual name or type of cancer was not needed to answer the current research questions. Fifth, because the results of automatic stemming for the Dutch language was too rudimentary (i.e., resulting in uninterpretable and unreadable words), multiple words were grouped together manually by adding this information to the code (e.g., *hooii* and *hallo* became *hoi* [hi]). To limit the influence of the noninformative difference between multiple conjugations of verbs on the topics, verb conjugations were grouped by the verb stem.

3.3. Parameters

To run the LDA model, values for the number of expected topics and the *alpha* parameter needed to be determined. Values of 0.10, 0.05, 0.01, 0.005, and 0.001 were tried. In line with the recommendation of [Jacobi et al. \(2016\)](#), a low *alpha* value was used to obtain fewer topics per document (i.e., $\alpha = 0.001$).

To establish the number of topics, analyses were performed on 5, 10, 15, 20, 25, 30, and 40 topics. For a maximum of 25 topics, the meaning of each topic could be identified without too much overlap between topics. Next, consensus about the meaning of the 25 topics was reached during meetings with the research team.

3.4. Differences between platforms

To answer the second research question, for all 25 topics, linear equations from topic modeling were applied to the sample of posts. This process resulted in a score for all 25 topics per document that measured the extent to which these topics occurred within each document. Then, per platform mean topic scores were calculated for all 25 topics. Finally, MANOVA was used to gain insight into the differences in the mean topic scores between the platforms. The 25 topics were used as dependent variables and the platforms as independent variables. When the Levene's F test showed significant differences between the four platforms, Bonferroni post hoc tests were used to gain more insights into these differences.

4. Results

4.1. Data

The post sample size was 52,995. This sample of documents consisted of all blog posts and replies ($n = 22,662$), all forum entries ($n = 9573$), all expert questions and answers ($n = 16,458$), and all entries into the medical libraries ($n = 4338$). The average number of words per document for each platform was $M_{\text{blogs}} = 78.55$ ($SD = 159.89$), $M_{\text{forum}} = 53.03$ ($SD = 64.01$), $M_{\text{questions}} = 41.62$ ($SD = 42.26$), and $M_{\text{library}} = 163.72$ ($SD = 131.47$).

Table 2
Overview of subtopics.

Subtopic number	Subtopic name	Topwords	Subtopic description	Example post
1	Sharing current status	Hi, cure, chemo, first, day, feeling, head	Patients share their recent experience mostly about treatment.	Hi everyone, today I had the first day of my chemo cure. I'm feeling pretty well, only a little dizzy in my head. (Blog)
2	Sharing previous experience	Months, years, ago, type of cancer, chemo, since, surgery, after	Patients share their previous experiences with cancer.	I also had *type of cancer* and since/after I received chemo/surgery two months/years ago I try to get back to living. (Forum response)
3	Cancer-treatment information	Type of cancer, treatment(s), doctor, tumor, examination, complaint, talk	Information about the treatment of cancer.	*Type of cancer*: common complaints include tiredness, loss of appetite and weight loss. If you experience any of these complaints talk to your doctors for an examination. (Library post)
4	Emotions and cancer	Life, emotions, fear, feeling, children, partner, difficult, talk	The emotional struggle with cancer is discussed in relation to relatives.	Emotions and Cancer: it is common to experience feelings of fear when diagnosed with cancer. How difficult it might be, try talking to your children and partner about it. (Library post)
5	Member interaction	Name, greetings, hi, dear, reaction, thanks, reply, story, experience	Interaction between members and the exchange of experiences.	Hi/Greetings dear *Name*, thank you/thanks for your reply/story/experience. (Forum response)
6	Treatment side effects	Pain, bother, cure, skin, side effects, complain, fluid, blood	Information about side-effects of treatment.	Side effects: complaints often include, pain, blood loss, excess bodily fluids, irritated skin. Patients are also bothered with [...]. (Library post)
7	Affective support	Good, name, best of luck, hi, hope, greetings, think, luck, wish	Messages wishing someone good luck, showing affection and thinking about them.	Hi *name*, Good luck / best of luck. I hope everything works out. I think of you. Greetings, *name*. (Blog response)
8	Sharing cancer status relatives	Hospital, father, mother, type of cancer, scan, appointment, doctor, conversation	Sharing status of cancer experience of relatives.	Today I went with my mother/father with *type of cancer* to the hospital for a scan/ an appointment with a doctor. (Blog)
9	Cancer tests	Examination, surgery, MRI, tissue, scan, metastasis, home, physician, surgeon	Tests and treatment methods of cancer.	After an MRI/ Tissue scan / Scan for metastasis, you can go home to wait for the results. (Library post)
10	Rehabilitation	Bother, fatigue, recovery, complaints, condition, exercise, sport, rehabilitation	Information on how to deal with tiredness and how to obtain a better condition / way of living.	Rehabilitation: patients are often bothered with various complaints concerning fatigue: the best way to recover is to slowly increase your condition by exercise/sport. (Library post)
11	Impact cancer on relatives	Wig, bald, deceased, somber, children, sister, grandmother, father	The impact of cancer (by close relatives) on oneself/others.	My grandmother/sister/child/father is feeling very somber about being bald due to chemo and having to wear a wig. (Blog post)
12	After care	Kankernl, group, discussion group, foundation, walk-in clinic, contact, website, organization	Information about the after care, and sources that provide after care.	More information can be found on the website of kanker.nl. Here you will find information about discussion groups, walk-in clinics, and contact information. (Library post)
13	Cancer organisations	Type of cancer, information, Dutch Cancer Society [kwf], Dutch Federation of Cancer Patients' Organizations [mfk], kankernl, website, colofon, patient association, livingwithcancer	Organizations and sources of information about cancer.	For more information on *type of cancer* the following websites of organizations can be contacted: Dutch Cancer Society [kwf], Dutch Federation of Cancer Patients' Organizations [mfk]; kanker.nl, [website, colofon], patient association, livingwithcancer [leven met kanker]. (Library post)
14	Cancer detection	Scan, arm, lymphedema, cea, contrast fluid, pet, blood, microscope, value, substance	Ways in which cancer will be tested.	Pet scan, blood inspection under a microscope, a test of your cea value, or contrast fluids are common ways to test for cancer. (Arm) lymphedema is an often-mentioned side effect of treatment. (library post)
15	Insurance/financial	Reimbursed, (life)insurance, money, payment, own, cost, euro, experimental, practical support	Information about insurance and coverage of different treatments and facilitating services.	Insurance will reimburse most (non-experimental) treatment. However, some costs for example for practical support (up to xxx euro) need to be paid for by yourself. (Library post)
16	Cancer research	Article, study, research, scientific, sexuality, curcuma, test, clinical, effect	Information about cancer research and related topics.	A new research article/ scientific study of a clinical test of the effects of Curcuma onl...]. (Library post)
17	Nutrition	Nutrition, food, sugar, fruit, meat, drink, tea, fish, fat, diet	Information about how to deal with diet and nutrition both during and after treatment.	Nutrition: a good diet could help recovery, some foods and drinks could help: try to eat fruit and fish/meat. Drink lots of tea and avoid too much sugar. (Library post)
18	Employment	Work(ing), employer, Employee Insurance Agency [UWV], sick, insurance, company doctor, colleague, supervisor, contract	Information about employment, laws, and organizations concerned with work and disease.	How to deal with working during cancer? If you have or had a contract with your employer possibly has insurance for sick leave. Talk to your employer and company doctor to discuss the options. In general, you are covered by the Employee Insurance Agency and you should contact them too. (Library post)
19	Other (alternative) treatments	Cannabis, weed oil, usage, oxygen therapy, cold cap	Information about treatments to limit the side effects or to address cancer itself.	

(continued on next page)

Table 2 (continued)

Subtopic number	Subtopic name	Topwords	Subtopic description	Example post
20	Medical trials	Trial, reimbursement, guinea pig, gene, database, patient card	Information about medical trials and costs.	To deal with side effects some patients report positive experiences with the use of a cold cap after radiation, oxygen therapy or the use of cannabis/weed oil. (Library post)
21	Gastrointestinal treatment	Stoma, surgery, bowel, colon, gist, stool, oxaliplatin, hipec	Colorectal cancer specific information.	A medical trial database exists in which for example gene therapy and reimbursement is discussed. (Library post)
22	Hereditary	Mutation, hereditary, condition, genetic testing, pregnancy	Information about the hereditary of cancer including tests and pregnancy.	For colorectal cancer or gist surgery and oxaliplatin-based hipec in combination with a stoma are treatment options. (Library post)
23	Treatment consequences	Bladder, incontinence, neuropathie, lipofilling, retained, spasm, catheter, prostate, prosthesis	Information about living with the consequence of treatment.	Questions about the genetic and hereditary side of cancer often come to mind during pregnancy. Genetic testing could help reduce stress and shed light on mutations. (Library post)
24	Information sources	Library, rehabilitation program, to inform, look, placed, movie	Discussion of information recourses.	After treatment, especially when the prostate is involved, bladder problems and incontinence are a problem. (Library post)
25	External organizations	Care, cooperation, Davinci, Dutch Cancer Society [kwf], organization, cancer research, alpe d'HuZes	Organizations that are focused on research, development, or appliance of cancer medication or raising money to do so.	More information can be found by taking a look in the library of this website, following a rehabilitation program or watching this movie to inform you about this topic. (Library post)

4.2. Topic identification

We wanted to gain insight into which social support topics could be distinguished on different cancer-related platforms of a health website and the extent to which these subtopics fit within the model of social support needs (RQ1). Many of the subtopics identified are similar to the social support topics that are part of the model of patients' social support needs. This similarity means many support needs identified within previous literature on offline information sources exist online. Fig. 1 provides a visual representation of the subtopics. Furthermore, the percentages displayed under Fig. 1 show the relative proportion of a subtopic in the complete sample of posts (i.e., the dominance of a certain subtopic within the content analyzed). Furthermore, Fig. 1 shows several (partly) overlapping subtopics, or subtopics that were similar to each other (i.e., certain words belong to different subtopics). In Table 2, the integration of the subtopics into the model is summarized, and example posts of each subtopic are given. We elaborate on this process below per layer of the social support model. Which Fyru Furthermore, we discuss the extent to which the predefined topics can be identified in the results of our topic model.

4.2.1. Layer II: social support category—informational support

Based on our theoretical model we expected to find topics ranging from cancer-specific information to tangible assistance that could be grouped under this support category (see Table 3). The analysis revealed that several subtopics were grouped under the social support topic *cancer-specific information*, such as cancer tests (subtopic 9) and cancer detection (subtopic 14). Additionally, several subtopics were part of the social support topic *treatment-related information*: cancer-treatment information (subtopic 3), treatment side effects (subtopic 6), other (alternative) treatments (subtopic 19), gastrointestinal treatment (subtopic 21), and treatment consequences (subtopic 23). Thus, when people talk about cancer-specific information, they often use words related to for example detection such as lymphedema, carcinoembryonic antigen test [CEA], contrast fluid, etc. When people talk online about treatment-related information, they use words related to cancer treatment information such as type of cancer or treatment(s).

Subtopics covering rehabilitation were found and grouped under the social support topic *rehabilitation information*. These subtopics are rehabilitation (subtopic 10) and after care (subtopic 12). In addition, subtopics covering the social support topic *surveillance and health information* included nutrition (subtopic 17) and hereditary (subtopic 22). We found that the subtopics emotions and cancer (subtopic 4) and impact cancer (relatives) (subtopic 11) were grouped under the social support topic *interpersonal/social information*. The social support topic *financial/legal information* included the subtopics insurance/financial (subtopic 15) and employment (subtopic 18). To illustrate, content about rehabilitation information features words like fatigue, recovery, and complaints. However, content included in the subtopic interpersonal/social information featured words such as life, emotions, fear, and feeling.

Within the social support topic *medical system information*, the following subtopics were found: cancer organizations (subtopic 13), cancer research (subtopic 16), and medical trials (subtopic 20). The social support topic *referral to other information sources* included the topics: information sources (subtopic 24) and external organizations (subtopic 25). In summary, the *medical system information* contained content around the words such as Dutch Cancer Society [kwf], trail, research, and scientific.

Based on our theoretical model we expected the presence of the social support topics *tangible assistance*, *prognosis information*, and *body image/sexuality information*. However, the results of our topic model included no such subtopics.

4.2.2. Layer II: social support—esteem, network and emotional support

Also for this layer, we investigate the extent to which and how the

Table 3
Mean topic scores comparison per platform.

Social support category	Social support topic	#	Subtopic	Blog ^a	Forum ^b	Library ^c	Questions ^d	
Informational support	Cancer-specific information	9	Cancer tests	2.38 ^{bcd}	1.67 ^{ac}	13.95 ^{abd}	1.83 ^{ac}	
		14	Cancer detection	1.26 ^{bcd}	0.57 ^{ac}	7.73 ^{abd}	0.44 ^{ac}	
	Treatment-related information	3	Cancer-treatment information	2.58 ^{bcd}	4.56 ^{ac}	50.53 ^{abd}	4.32 ^{ac}	
		6	Treatment side effects	3.51 ^{bcd}	2.45 ^{ac}	9.73 ^{abd}	2.46 ^{bc}	
		19	Other (alternative) treatments	0.53 ^{bc}	0.68 ^{acd}	1.97 ^{abd}	0.49 ^{bc}	
		21	Colorectal cancer	0.48 ^c	0.43 ^c	2.50 ^{abd}	0.41 ^c	
		23	Treatment consequences	0.61 ^{bc**d}	0.47 ^{ac}	0.74 ^{a**bd}	0.46 ^{ac}	
	Rehabilitation information	10	Rehabilitation	2.11 ^{cd}	2.09 ^{cd**}	14.60 ^{abd}	1.55 ^{ab**c}	
		12	After care	1.33 ^{bcd}	1.77 ^{acd}	8.66 ^{abd}	0.73 ^{abc}	
	Surveillance and health information	17	Nutrition	0.90 ^{cd}	0.70 ^c	2.85 ^{abd}	0.54 ^{ac}	
		22	Hereditary	0.65 ^{bcd}	0.41 ^{acd}	1.84 ^{abd}	0.29 ^{abc}	
	Interpersonal / social information	11	Impact cancer (on relatives)	4.98 ^{bcd}	1.22 ^{ad*}	0.95 ^a	0.72 ^{ab*}	
			Emotions and cancer	6.80 ^{bcd}	4.32 ^{acd}	18.08 ^{abd}	2.11 ^{abc}	
	Financial / legal information	15	Insurance/financial	1.40 ^{bcd}	0.97 ^{acd}	1.84 ^{abd}	0.50 ^{abc}	
		18	Employment	1.00 ^{bcd}	0.74 ^{acd}	2.43 ^{abd}	0.40 ^{abc}	
	Medical system information	13	Cancer organisations	0.60 ^{bc}	0.95 ^{acd}	10.42 ^{abd}	0.57 ^{bc}	
			Cancer research	0.97 ^{cd}	1.14 ^{cd**}	3.33 ^{abd}	0.56 ^{ab**c}	
		20	Medical trails	0.78 ^{bcd}	0.44 ^{acd**}	1.55 ^{abd}	0.31 ^{ab**c}	
		24	Information sources	0.40 ^{bcd}	0.29 ^{ac}	0.97 ^{abd}	0.23 ^{ac}	
	Referral other information sources	25	External organisations	0.36 ^{bcd}	0.21 ^{acd*}	0.90 ^{abd}	0.16 ^{abc}	
		7	Affective support	4.97 ^{bcd}	2.70 ^{acd**}	0.34 ^{abd}	2.40 ^{ab**c}	
	Esteem support	Coping and spiritual	7	Affective support	4.97 ^{bcd}	2.70 ^{acd**}	0.34 ^{abd}	2.40 ^{ab**c}
	Network support	Support from others and groups	5	Member interaction	5.28 ^{bcd}	6.07 ^{acd}	1.12 ^{abd}	5.58 ^{abc}
	Emotional support	Sharing status	1	Sharing current status	19.55 ^{bcd**}	5.05 ^{acd}	1.40 ^{ab}	3.05 ^{ab}
			2	Sharing previous status	10.20 ^{cd}	10.51 ^{cd}	2.79 ^{abd}	9.48 ^{abc}
8			Sharing status relatives	4.85 ^{bcd}	2.64 ^{ad}	2.52 ^a	2.02 ^{ab}	

Note. All letters, unless otherwise specified, indicate $p < .001$. For example: ^{bcd} indicates the mean score for that topic differ significantly from the mean score for that topic on the platforms forum, library, and questions with $p < 0.001$.

* $p < .05$. ** $p < .01$.

predefined topics from the social support model could be found in the outcome of the topic model. First, the social support topic *coping and spiritual* was found and included the subtopic affective support (subtopic 7). Furthermore, *support from others and groups* was found and included subtopic member interactions (subtopic 5). The social support topic *sharing status* was found in the following subtopics: sharing current status (subtopic 1), sharing previous experience (subtopic 2), and sharing status relatives (subtopic 8). Based on the visualization of the topics (see Fig. 1) these subtopics are in the proximity of the affective subtopic, indicating a close semantic resemblance (i.e., subtopics contain similar top words). Here, users proactively posted status updates about their current status and reactively posted their previous status or on the status of relatives. Because the subtopics are better grouped under the social support topic named *sharing status* than under the name *emotional support*, the label was changed to *sharing status* (see Table 3).

4.2.3. Different platforms for different needs

We explored how peer- and expert-generated cancer-related platforms differ from each other in terms of the social support provided. Overall, the MANOVA shows significant differences between the platforms of all 25 subtopics (see Table 3 for the complete table). While the social support topics were not strictly divided between the platforms, the platforms differed in the degree to which social support topics were present in the different platforms.

5. Informational support

Informational support subtopics were found most in the library platform of the website (9, 14, 3, 6, 19, 21, 23, 10, 12, 17, 22, 4, 15, 18, 13, 16, 20, 24, and 25). To illustrate, within the topic *cancer specific information*, the subtopic cancer tests (subtopic 9) is covered more in the library than it is in all other platforms, including expert questions ($M_{\text{library}} = 13.95$, $M_{\text{blog}} = 2.38^{*1}$, $M_{\text{forum}} = 1.67^{*}$, and $M_{\text{Questions}} = 1.83^{*}$, see Table 3). Within the topic *treatment-related*

information, the subtopic cancer treatment information is most dominant, i.e., this subtopic was covered more often in the library than in all other platforms (see subtopic 3 in Table 3). The same holds true for the following topics within informational support: *rehabilitation information* (rehabilitation; subtopic 10), *surveillance and health information* (nutrition; subtopic 17), *financial and legal information* (employment; subtopic 18), *medical system information* (cancer organizations; subtopic 13), and *referral to other sources* (information sources; subtopic 24). The topic *interpersonal/social information* showed mixed findings. While emotions and cancer (subtopic 4) were covered more within the library than they were in all other platforms, impact cancer (on relatives) (subtopic 11) was covered more on blogs than it was in all other platforms. In sum, the library showed the highest coverage in topics concerning general cancer-related information regarding the cancer patients' whole journey, from diagnosis through after care.

5.1. Esteem, network, and emotional support

Within the social support topic *coping and spiritual*, the subtopic affective support (subtopic 7) is prominent. This topic was covered more on blogs than it was on all other platforms ($M_{\text{blog}} = 4.97$, $M_{\text{forum}} = 2.70^{*}$, $M_{\text{library}} = 0.34^{*}$, and $M_{\text{Questions}} = 2.40^{*}$). Within the topic *support from others and groups*, member interaction (subtopic 5) was found, which was covered more on forums than it was on all other platforms. For *sharing status*, a nuanced image appears: sharing current status (subtopic 1) and sharing status relatives (subtopic 8) were covered more on blogs than they were on discussion forums and other platforms ($M_{\text{blog_subtopic1}} = 19.28$, $M_{\text{forum_subtopic1}} = 5.05^{*}$, $M_{\text{library_subtopic1}} = 1.40^{*}$, $M_{\text{Questions_subtopic1}} = 3.05^{**2}$; $M_{\text{blog_subtopic8}} = 4.85$, $M_{\text{forum_subtopic8}} = 2.64^{*}$, $M_{\text{library_subtopic8}} = 2.52^{*}$, and $M_{\text{Questions_subtopic8}} = 2.02^{*}$). Sharing previous experience, however, was found equally on blogs and discussion forums, where it was found more often than it was on the library and question-and-answers platforms.

¹ $p < .001$

² $p < .05$

6. Discussion

This study investigated the different social support needs and topics on online health information platforms. A theoretical model integrating social support needs and topics was developed (i.e., the model of patients' social support needs) and applied to different online platforms using computer-assisted content analysis. Results showed differences between expert- and peer-generated platforms and differences within these types of platforms. Expert-generated platforms generally included informational support and related social support topics focusing on cognitive needs, whereas peer-generated platforms generally included affective support and related social support topics focusing on affective needs. When looking at the differences within one type of platform we found that users talk differently about their experiences on different peer-generated platforms (blogs and forums). These results demonstrate that the platforms fundamentally differ in terms of content. As these platforms may influence the fulfillment of social support needs we can expect that these platforms also affect patient outcomes differently. Future research should therefore explicitly distinguish which platforms are being studied.

On a methodological level, the current study showed the benefits of using a hybrid method, in which a computational social science method is combined with a theoretical a priori-defined model. It allowed us to use a theoretical model to aid interpretation of the results but at the same time to remain open for unexpected findings. For example, many of the subtopics identified were similar to the social support topics of the theoretical model. However, the results indicated that the original model should be extended with the additional social support topic named sharing status. To provide informational support, users shared (previous) status updates about themselves or relatives with other users. Although not included in previous typologies of patient needs, previous research highlighted the patient's need to provide their knowledge, quote their experiences, and share their progress online (Eysenbach et al., 2004; Gooden and Winefield, 2007; Coulson, 2005; Wang et al., 2015). If a supervised machine learning approach was used instead of an unsupervised machine learning approach this topic would have remained missing in the proposed model. In sum, the current approach allowed for additional topics to emerge from the data.

While most of the social support topics as identified in the theoretical model were found, three topics were not identified (i.e., prognosis information, body image, and tangible assistance). This finding does not necessarily mean that these topics were not covered in the online context per se. We chose to conduct our analysis with 25 topics. While 25 topics made interpretation of the topics possible, this number might not have been sufficiently discriminative to differentiate between very specific topics. Another explanation can be found in the definition used. For example, the definition used to identify tangible assistance was maybe difficult to find online (Coulson et al., 2007). Words in this definition such as "concrete" and "physical action" were previously mentioned concerning offline tangible assistance (Coulson et al., 2007). Future research should redefine the definition appropriate to the online context. The definition should cover examples such as "I'll email you details of this diet ... I hope it helps!" or "I am happy to help any way I can ... just let me know" (Coulson, 2005, p.583). While most of these examples do not provide tangible evidence of support, the willingness to help is evident in this definition.

To further specify and expand the proposed model, the current study should be replicated using different analyses and perspectives. As an illustration, our model can be used as a starting point for future research for supervised machine learning on different online platforms. By applying this method, a more nuanced picture and more specific topics can be identified. For example, using annotated data on how certain topics are discussed would offer a deeper understanding of the nature of patients' needs. While the current study included multiple, highly-used platforms, these platforms were all part of one source, a Dutch website, that focused on one illness, i.e., cancer. Therefore,

generalizability to other countries, platforms, illnesses, and cultures is unknown. For instance, popular social media platforms such as Facebook, Twitter, and YouTube contain "countless discussions of health, illness, and medicine" (Lupton, 2016, p.55). These platforms offer different experiences to the user (e.g., more direct contact, closer-knit groups, and videos) and receive scarce scholarly attention. Analyzing these platforms using the current theoretical model could replicate and validate our findings as well as further improve our understanding of online health information platforms. The model can be further expanded by including more types of diseases across more countries and cultures. Especially when the diseases are considered taboo. In this latter case, online health information might be even more important to these patients. Similar differences in reliance on different sources of health information could be found in other countries and cultures in which medical experts' expertise is more highly valued. Possibly this limits the dependence on online health information. Future research in other countries showing a higher difference in power between individuals (e.g., China, Japan, and Brazil; Hofstede, 2011), and thus more reliance on experts for information could broaden our understanding of the role online health information in the context of other sources.

Last, researchers are encouraged to consider group interaction on blogs and discussion forums. Due to the current individual focus in our model, the kinds of needs fulfilled by the online interaction between users remain unclear. Given the large degree to which these responses are given, clearly, some needs are fulfilled. Furthermore, due to the nature of the analysis applied, only the needs and content of posters are analyzed. However, previous research showed most users do not actively write on these platforms but merely read them (i.e., lurkers; Coulson, 2005). Future research should therefore apply different techniques (e.g., pop-up online surveys or focus groups) to identify the needs and motives of this group of users.

Technological advances are changing the way patients search for information, manage their health and the patient-provider relationship (Asan and Carayon, 2017). Technological and health care advances provide the means to more effectively and less expansively manage chronic diseases (Stephanidis et al., 2019). With these technological advances and the availability of online health information platforms, more content (i.e., data) is created (Shah et al., 2015; Topol, 2015; Sillence et al., 2006). How this data is gathered, shared and, used can further improve our understanding of how different stakeholders use this information to manage health (Victorelli et al., 2019; Branford, 2019). As a result of this ever-increasing amount of data, the need to better order and search content online becomes a priority. Insights originating from the current study could influence practice in this regard. Two main suggestions that will help aid and improve content ordering, patients' health information-seeking behavior, and in turn, patient-provider interaction will be discussed. On the one hand, design factors in the interface could help users to navigate and match content to their needs. The model of patients' social support needs can be used as a blueprint for ordering online content. By ordering the content by the proposed model, the content will follow the same route from very "abstract" social support needs towards increasingly precise needs and topics (Wilson et al., 2002). Second, using the algorithms to annotate the data based on patients' social support needs, the content can automatically be accompanied by visuals that indicate to the users which topic is discussed. For example, platforms can include visual cues such as an icon of a brain for cognitive content or an icon of a heart for affective content. Alternatively, parts of texts that match the social support needs of the user can be highlighted or include an exclamation mark that would signal patients to the relevancy of those parts based on their indicated social support needs. This enables patients to intuitively browse and more easily find content based on their needs. A well-designed website, with a focus on information seeking (ordering) and integration (visual cues), could benefit patients by limiting the possibility of encountering misinformation and/or reducing cognitive

overload (Lyu and Li, 2019; Blandford et al., 2014).

On the other hand, our results can help providers, patients, and relatives to find content more easily and in turn, positively influence patient-provider interaction and patient outcomes. The algorithms used in our analysis can be used to optimize searches from a keyword-based approach to a semantic similarity-based approach in which the context of keywords are taken into account. The word clusters identified during our analysis included words that are linked together through their latent meaning and thus provide context to each other. These can be used to help users find the right content. Search approaches like semantic similarity, are found to provide better search results on platforms such as forums (Franco-Salvador et al., 2018). The context of the words will help with overcoming problems due to polysemy (i.e., words with multiple meanings depending on the context) and synonymy words (Malve and Chawan, 2015). For example, using a semantic similarity-based approach instead of a keyword-based approach, a patient searching with a search query like “weed oil” or “alternative treatment” will not only see content that include that specific keyword, but also parts that contain synonyms or related complementary treatments (e.g., the oil, complementary treatment, or acupuncture). These words are part of the same word cluster and thus might fulfill the need of a user searching for alternative treatments. This will make a patients’ search strategy more intuitive and complete. In the future, we could formalize these semantics into ontologies of which the proposed model could be a starting point. This ontology would even allow for searching across multiple sources of information (Malve and Chawan, 2015). In sum, these two applications have the potential to make online content more accessible, easier to navigate, and tailored towards patients’ needs. In a health care setting where patients’ empowerment is important but the patients also deal with emotional and cognitive uncertainty due to their health condition, easy access to health information is an important topic (Huang and Haried, 2019). A more intuitive and better functioning search engine will make it easier to find more relevant and complete information particularly for older patients and those with low health or technology literacy (Blandford, 2019; Topol, 2015; Becker, 2004). Research shows that patients who are more confident about their search results are more likely to engage in talking about that information with their medical expert (Diaz et al., 2002; Sommerhalder et al., 2009). These discussions in turn allow for an appraisal of the online information, resolving misunderstandings, a higher degree of shared decision making, and better patient outcomes (Benetoli et al., 2018; Sanders and Linn, 2018; Asan and Carayon, 2017).

Declaration of Competing Interest

None.

References

- Abdul-Muhsin, H., Tyson, M., Raghu, S., Humphreys, M., 2015. The informed patient: an analysis of information seeking behavior and surgical outcomes among men with benign prostatic hyperplasia. *Am. J. Men's Health*. <https://doi.org/10.1177/1557988315617527>.
- Anderson, J.G., Rainey, M.R., Eysenbach, G., 2003. The impact of cyberhealthcare on the physician-patient relationship. *J. Med. Sys.* 27 (1), 67–84. <https://doi.org/10.1023/A:1021061229743>.
- Asan, O., Carayon, P., 2017. Human factors of health information technology—challenges and opportunities. *Int. J. Hum.-Comput. Interact.* 33 (4), 255–257. <https://doi.org/10.1080/10447318.2017.1282755>.
- Beck, U., Giddens, A., Lash, S., 1994. *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Stanford University Press.
- Becker, S.A., 2004. A study of web usability for older adults seeking online health resources. *ACM Trans. Comput.-Hum. Interact.* 11 (4), 387–406. <https://doi.org/10.1145/1035575.1035578>.
- Benetoli, A., Chen, T.F., Aslani, P., 2018. How patients’ use of social media impacts their interactions with healthcare professionals. *Patient Educ. Couns.* 101 (3), 439–444. <https://doi.org/10.1016/j.pec.2017.08.015>.
- Bensing, J., Verhaak, P., 2004. Communication in medical encounters. In: Weinman, J., Kaptain, A. (Eds.), *Health Psychology*, Eds. Blackwell Publishing, Malden, pp. 261–287.
- Blandford, A., 2019. HCI for health and wellbeing: challenges and opportunities. *Int. J. Hum.-Comput. Stud.* 131, 41–51. <https://doi.org/10.1016/j.ijhcs.2019.06.007>.
- Blandford, A., Furniss, D., Vincent, C., 2014. Patient safety and interactive medical devices: realigning work as imagined and work as done. *Clin. Risk* 20 (5), 107–110. <https://doi.org/10.1177/1356262214556550>.
- Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, 993–1022. <https://doi.org/10.1162/jmlr.2003.3.4.5.993>.
- Buis, L.R., Whitten, P., 2011. Comparison of social support content within online communities for high- and low-survival-rate cancers. *Comput. Inform. Nurs.* 29 (8), 461–467. <https://doi.org/10.1097/NCN.0b013e318214093b>.
- Cotten, S.R., Gupta, S.S., 2004. Characteristics of online and offline health information seekers and factors that discriminate between them. *Soc. Sci. Med.* 59 (9), 1795–1806. <https://doi.org/10.1016/j.socscimed.2004.02.020>.
- Coulson, N.S., 2005. Receiving social support online: An analysis of a computer-mediated support group for individuals living with irritable bowel syndrome. *Cyber Psychol. Behav.* 8 (6), 580–584. <https://doi.org/10.1089/cpb.2005.8.580>.
- Coulson, N.S., Buchanan, H., Aubeeluck, A., 2007. Social support in cyberspace: a content analysis of communication within a huntington’s disease online support group. *Patient Educ. Couns.* 68 (2), 173–178. <http://dx.doi.org/10.1016/j.pec.2007.06.002>.
- Cutrona, C.E., 1990. Stress and social support – in search of optimal matching. *J. Soc. Clin. Psychol.* 9 (1), 3–14. <https://doi.org/10.1521/jscp.1990.9.1.3>.
- Cutrona, C.E., Russell, D.W., 1990. *Type of Social Support and Specific Stress: Toward a Theory of Optimal Matching*. John Wiley & Sons, Oxford.
- Cutrona, C.E., Suhr, J.A., 1992. Controllability of stressful events and satisfaction with spouse support behaviors. *Commun. Res.* 19 (2), 154–174. <https://doi.org/10.1177/009365092019002002>.
- Diaz, J.A., Griffith, R.A., Ng, J.J., Reinert, S.E., Friedmann, P.D., Moulton, A.W., 2002. Patients’ use of the Internet for medical information. *J. Gen. Intern. Med.* 17 (3), 180–185. <https://doi.org/10.1046/j.1525-1497.2002.10603.x>.
- Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., Stern, A., 2004. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Br. Med. J.* 328 (7449), 1166. <https://doi.org/10.1136/bmj.328.7449.1166>.
- Franco-Salvador, M., Kar, S., Solorio, T., & Rosso, P. (2018). UH-PRHLT at SemEval-2016 Task 3: combining lexical and semantic-based features for community question answering. arXiv preprint arXiv:1807.11584.
- Gooden, R.J., Winefield, H.R., 2007. Breast and prostate cancer online discussion boards: a thematic analysis of gender differences and similarities. *J. Health Psychol.* 12 (1), 103–114. <https://doi.org/10.1177/1359105307071744>.
- Goldsmith, D.J., 2004. *Communicating Social Support*. Cambridge University Press.
- Gutierrez, N., Kindratt, T.B., Pagels, P., Foster, B., Gimpel, N.E., 2014. Health literacy, health information seeking behaviors and internet use among patients attending a private and public clinic in the same geographic area. *J. Community Health* 39 (1), 83–89. <https://doi.org/10.1007/s10900-013-9742-5>.
- Hether, H.J., Murphy, S.T., Valente, T.W., 2014. It’s better to give than to receive: the role of social support, trust, and participation on health-related social networking sites. *J. Health Commun.* 19 (12), 1424–1439. <https://doi.org/10.1080/10810730.2014.894596>.
- Hofstede, G., 2011. Dimensionalizing cultures: the Hofstede model in context. *Online Read. Psychol. Cult.* 2 (1), 8. <https://doi.org/10.9707/2307-0919.1014>.
- House, J.S., 1983. *Work Stress and Social Support*. Addison-Wesley Series on Occupational Stress.
- Huang, C.L., Haried, P., 2019. An evaluation of uncertainty and anticipatory anxiety impacts on technology use. *Int. J. Hum.-Comput. Interact.* 1–9. <https://doi.org/10.1080/10447318.2019.1672410>.
- Iconomou, G., Vagenakis, A.G., Kalofonos, H.P., 2001. The informational needs, satisfaction with communication, and psychological status of primary caregivers of cancer patients receiving chemotherapy. *Support. Care Cancer* 9 (8), 591–596. <https://doi.org/10.1007/s005200100259>.
- Jacobi, C., van Atteveldt, W., Welbers, K., 2016. Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digit. Journal.* 4 (1), 89–106. <https://doi.org/10.1080/21670811.2015.1093271>.
- Klemm, P., Bunnell, D., Cullen, M., Soneji, R., Gibbons, P., Holecsek, A., 2003. Online cancer support groups: a review of the research literature. *Comput. Inform. Nurs.* 21 (3), 136–142. <https://doi.org/10.1097/00024665-200305000-00010>. 10.2196/jmir.7.1.e10.
- Leydon, G.M., Boulton, M., Moynihan, C., Jones, A., Mossman, J., Boudioni, M., McPherson, K., 2000. Cancer patients’ information needs and information seeking behaviour: in depth interview study. *Br. Med. J.* 320 (7239), 909. <https://doi.org/10.1136/bmj.320.7239.909>.
- Lupton, D., 1997. Consumerism, reflexivity and the medical encounter. *Soc. Sci. Med.* 45 (3), 373–381. [https://doi.org/10.1016/s0277-9536\(96\)00353-x](https://doi.org/10.1016/s0277-9536(96)00353-x).
- Lupton, D., 2013. The digitally engaged patient: self-monitoring and self-care in the digital health era. *Soc. Theory Health* 11 (3), 256–270. <https://doi.org/10.1057/sth.2013.10>.
- Lupton, D., 2016. Towards critical digital health studies: Reflections on two decades of research in health and the way forward. *Health* 20 (1), 49–61. <https://doi.org/10.1177/1363459315611940>.
- Lyu, X., Li, Z., 2019. Correlations between human performance in information seeking, information integration, and overall process in diagnostic tasks. *Int. J. Hum.-Comput. Interact.* 1–10. <https://doi.org/10.1080/10447318.2019.1631545>.
- Malve, A., Chawan, P.M., 2015. A comparative study of keyword and semantic based search engine. *Int. J. Innov. Res. Sci. Eng. Technol.* 4 (11), 11156–11161. <https://doi.org/10.15680/IJRSET.2015.0411039>.
- McCombs, M.E., Shaw, D.L., 1972. The agenda-setting function of mass media. *Public*

- Opin. Q. 36 (2), 176–187. <https://doi.org/10.1086/267990>.
- Namkoong, K., McLaughlin, B., Yoo, W., Hull, S.J., Shah, D.V., Kim, S.C., ... Gustafson, D.H., 2013. The effects of expression: how providing emotional support online improves cancer patients' coping strategies. *J. Natl. Cancer Inst.* 2013 (47), 169–174. <https://doi.org/10.1093/jncimonographs/igt033>.
- Niederdeppe, J., Hornik, R.C., Kelly, B.J., Frosch, D.L., Romantan, A., Stevens, R.S., ... Schwartz, J.S., 2007. Examining the dimensions of cancer-related Information seeking and scanning behavior. *Health Commun.* 22 (2), 153–167. <https://doi.org/10.1080/10410230701454189>.
- Petz, G., Karpowicz, M., Fürschuß, H., Auinger, A., Střiteský, V., Holzinger, A., 2014. Computational approaches for mining user's opinions on the Web 2.0. *Inf. Process. Manag.* 50 (6), 899–908. <https://doi.org/10.1016/j.ipm.2014.07.005>.
- Prestin, A., Vieux, S.N., Chou, W., 2015. Is online health activity alive and well or flattening? Findings from 10 years of the health information national trends survey. *J. Health Commun.* 20 (7), 790–798. <https://doi.org/10.1080/10810730.2015.1018590>.
- Puts, M.T.E., Papoutsis, A., Springall, E., Tourangeau, A.E., 2012. A systematic review of unmet needs of newly diagnosed older cancer patients undergoing active cancer treatment. *Support. Care Cancer* 20 (7), 1377–1394. <https://doi.org/10.1007/s00520-012-1450-7>.
- Rains, S.A., Young, V., 2009. A meta-analysis of research on formal computer-mediated support groups: examining group characteristics and health outcomes. *Hum. Commun. Res.* 35, 309–336. <https://doi.org/10.1111/j.1468-2958.2009.01353.x>.
- Řehůřek, R., Sojka, P., 2010. Software framework for topic modelling with large corpora. In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Valletta, Malta. ELRA, pp. 45–50.
- Rupert, D.J., Read, J.G., Amoozegar, J.B., Moultrie, R.R., Taylor, O.M., O'Donoghue, A.C., Sullivan, H.W., 2016. Peer-generated health information: the role of online communities in patient and caregiver health decisions. *J. Health Commun.* 21 (11), 1–11. <https://doi.org/10.1080/10810730.2016.1237592>.
- Rutten, L.J.F., Arora, N.K., Bakos, A.D., Aziz, N., Rowland, J., 2005. Information needs and sources of information among cancer patients: a systematic review of research (1980–2003). *Patient Educ. Couns.* 57 (3), 250–261. <https://doi.org/10.1016/j.pec.2004.06.006>.
- Sanders, R., Linn, A.J., 2018. A mixed method study investigating the impact of talking about patients' internet use on patient-reported outcomes. *J. Health Commun.* 23 (9), 815–823. <https://doi.org/10.1080/10810730.2018.1514443>.
- Sanson-Fisher, R., Girgis, A., Boyes, A., Bonevski, B., Burton, L., Cook, P., 2000. The unmet supportive care needs of patients with cancer. Supportive care review group. *Cancer* 88 (1), 226–237. <https://doi.org/10.4103/0973-1075.138400>.
- Shah, D.V., Cappella, J.N., Neuman, W.R., 2015. Big data, digital media, and computational social science: possibilities and perils. *Ann. Am. Acad. Polit. Soc. Sci.* 659 (1), 6–13. <https://doi.org/10.1177/0002716215572084>.
- Sillence, E., Briggs, P., Harris, P., Fishwick, L., 2006. A framework for understanding trust factors in web-based health advice. *Int. J. Hum.-Comput. Stud.* 64 (8), 697–713. <https://doi.org/10.1016/j.ijhcs.2006.02.007>.
- Sommerhalder, K., Abraham, A., Zufferey, M.C., Barth, J., Abel, T., 2009. Internet information and medical consultations: experiences from patients' and physicians' perspectives. *Patient Educ. Couns.* 77 (2), 266–271. <https://doi.org/10.1016/j.pec.2009.03.028>.
- Stephanidis, C., Salvendy, G., Antona, M., Chen, J.Y., Dong, J., Duffy, V.G., ... Guo, Y., 2019. Seven HCI grand challenges. *Int. J. Hum.-Comput. Interact.* 35 (14), 1229–1269. <https://doi.org/10.1080/10447318.2019.1619259>.
- Topol, E., 2015. *The Patient Will See You Now: The Future of Medicine is in Your Hands*. Basic Books, New York.
- Victorelli, E.Z., Dos Reis, J.C., Hornung, H., Prado, A.B., 2019. Understanding human-data interaction: literature review and recommendations for design. *Int. J. Hum.-Comput. Stud.* 134, 13–32. <https://doi.org/10.1016/j.ijhcs.2019.09.004>.
- Wang, Y.-C., Kraut, R.E., Levine, J.M., 2015. Eliciting and receiving online support: using computer-aided content analysis to examine the dynamics of online social support. *J. Med. Internet Res.* 17 (4), e99. <https://doi.org/10.2196/jmir.3558>.
- Wilson, T.D., Ford, N.J., Ellis, D., Foster, A.E., Spink, A., 2002. Information seeking and mediated searching: part 2. Uncertainty and its correlates. *J. Am. Soc. Inf. Sci. Technol.* 53 (9), 704–715. <https://doi.org/10.1002/asi.10082>.
- Xiao, N., Sharman, R., Rao, H.R., Upadhyaya, S., 2014. Factors influencing online health information search: an empirical analysis of a national cancer-related survey. *Decis. Support Syst.* 57, 417–427. <https://doi.org/10.1016/j.dss.2012.10.047>.