DisVis: Visualizing Discussion Threads in Online Health Communities

Drashko Nakikj, MS, MA, Lena Mamykina, PhD
Department of Biomedical Informatics, Columbia University, New York, NY, USA

Abstract

An increasing number of individuals turn to online health communities (OHC) for information, advice and support about their health condition or disease. As a result of users' active participation, these forums store overwhelming volumes of information, which can make access to this information challenging and frustrating. To help overcome this problem we designed a discussion visualization tool DisVis. DisVis includes features for overviewing, browsing and finding particular information in a discussion. In a between subjects study, we tested the impact of DisVis on individuals' ability to provide an overview of a discussion, find topics of interest and summarize opinions. The study showed that after using the tool, the accuracy of participants' answers increased by 68% (p-value = 0.023) while at the same time exhibiting trends for reducing the time to answer by 38% with no statistical significance (p-value = 0.082). Qualitative interviews showed general enthusiasm regarding tools for improving browsing and searching for information within discussion forums, suggested different usage scenarios, highlighted opportunities for improving the design of DisVis, and outlined new directions for visualizing user-generated content within OHCs.

Introduction

In recent years, online health communities (OHC) have emerged as an important complement to the traditional health care[1–3]. According to the 2009 Pew Report, close to a 1/3 of all American adults had accessed social media for questions related to health[4]. More than 1/2 of e-patients (patients relying on online medical services) living with chronic disease consume user-generated health information and 1/5 participate in generating that type of information[5]. In response, there emerged a large number of OHCs such as breastcancer.org, TuDiabetes.org, PatientsLikeMe.com, to name a few. Studies showed that these forums serve many purposes, from sharing information and increasing knowledge, to emotional and social support, to connecting with patients with similar health challenges, among others[1,2,6]. However, most of these forums continue to rely on text-based discussion boards as their main platform for achieving these goals, thus presenting rich repositories of text-based information[2,6,7]. Because of the amount of information collected within OHCs, satisfying users' information needs can be a complicated and time consuming process[6,8]. For example, based on the data collected in 2014, an active forum for diabetes self-management, TuDiabetes.org, has over 36,000 members and includes almost 23,000 discussions on broad range of topics. Many of these discussions include dozens of members and hundreds of posts, and last for several months or years, transforming the initial topic of the discussion in multiple directions[8]. Therefore, finding relevant information and forming an impression of the topics discussed and opinions expressed within a single discussion thread can be challenging. While in many domains this can create inefficiencies and lead to frustration, in the context of health information it can lead to particularly disastrous effects, such as participants arriving at wrong conclusions to the detriment to their health.

To address these issues, the long-term objective of this research is to develop novel discussion visualization tools that will help members of OHCs to navigate discussion threads, quickly form an impression of what is being discussed, and identify the most interesting and relevant parts of the conversation. In this paper we present the prototype of the DisVis discussion visualization tool and discuss the results of its evaluation study with individuals recruited from TuDiabetes. In this between subjects study, the participants in the intervention and the control group were asked to review a single discussion thread and were asked to generate an overview of discussion topics and participants, find topics of interest, and summarize opinions while using DisVis (experimental group) or the original discussion board representation (control group). We compared individuals' answers to a gold standard, generated by either members of the research team or a healthcare professional; the accuracy of individuals' answers and time to answer were captured and analyzed. The evaluation session concluded with short qualitative interviews assessing individuals' subjective impressions from the tool and opportunities for improvement. We will present the detailed results from this study and their implications for the design of new platforms for OHCs in the remainder of the paper.

Related Work

OHCs have become a topic of active investigation within several research communities, with the particular attention on information seeking practices of their members. Previous research within OHCs found that users expressed

higher levels of satisfaction with their experience when their information needs were matched with appropriate information[9]. Along these lines, previous studies explored common characteristics of health information seeking. These studies suggested that individuals prefer answers by others matched on profile similarity[10], highlighted the importance of context when posing queries, and suggested that search results should be personalized based on user's medical history[11]. Consequently, recent trends in the research on OHCs increasingly favor new computational solutions for optimizing access to information[12,13], expertise and interest matching[14,15], mining members' contributions over time to infer their credibility[16], and using automated ways to extract personal information from members' contributions[17,18]. In regards to information seeking in OHCs, Nambisan suggested the need to focus on developing tools that make information seeking more effective and efficient[19].

Beyond health-oriented discussion forums, other domains have explored the question of accessing discussion information through visualization in great depth, including visualization of enterprise flash forums, e-mails, and newsgroups. For example, ConVis[20] is a system that enables the user to quickly form a topic-post-sentiment-author chain connection when inspecting the discussion. ForumReader[21] is a tool developed for exploration and analysis of flash forums based on the thumbnail representation of the discussion. Themail[22] builds a visual display of e-mail based interactions over time, while Newsgroup Crowds[23] graphically represents the population of authors in a particular newsgroup in Usenet and AuthorLines[23] visualizes a particular author's posting activity in the same platform. Other researchers have looked at preserving and visualizing the tree structure in the discussions using the sunburst radial space filling approach[24], icicle tree layout[25] and adding reply-based sociogram view[26]. Other approaches to visualizing large amounts of information relevant for this research were explored within topic based visual text analysis[27], large network data visualization with rich user interaction[28] and visualizing social roles[29]. Finally, previous research explored ways to visualize opinions about a product[30] and opinions about ideas[31].

Despite this ongoing research on visualizing user-generated content collections, addressing information overload in OHC forums with visualization solutions has received limited attention in the past. Within this domain, previous research focused on topic modeling and network visualization to explore patients' experiences[12] and on modeling interpersonal interactions and medication use[13], exploring relationships between topics and emotions[32], as well as health behavior and symptoms polarity[33]. Another direction explored by previous research focused specifically on supporting community moderators. For example, VisOHC[34] is a tool for OHC moderators that integrally visualizes the forum, aggregating diverse dimensions of conversational threads.

Although these systems and tools represent a rich body of knowledge about forum/discussion visualization, few of them focused specifically on helping members of OHCs overview discussion threads within their community, find topics of interest, and synthesize different opinions on these topics. As a result, the goal of this work was to design a simple discussion visualization tool that was tailored to discussions within OHCs, and to evaluate it with members of one such community. In the spirit of technology probes[35], we intended this prototype as a tool for proving a concept: testing the success of basic design approaches and eliciting user feedback.

Methods

The Prototype: DisVis

The design of DisVis was informed by our own previous studies of collective sensemaking within TuDiabetes[8] and by review of related literature. Our particular focus was on assisting users in three main directions: 1) forming an overview of the entire discussion thread, 2) finding prevalent topics within a discussion, and 3) summarizing different opinions in regards to the prevalent topics. DisVis interface is logically divided into two main segments: *Exploration and Navigation area* (Figure 1-B) and *Full Content area with a summary* (Figure 1-A). These two segments were introduced to allow users to move seamlessly between exploring different aspects of the discussion and reading the text of interest in more detail, respectively. We describe these features in detail below.

Discussion overview. The first set of features described here was included to help individuals form a quick impression of the entire discussion thread. The *Summary* section (2) provides a brief summary of the discussion, including its key descriptors: number of posts, number of participants, duration, and keywords. The *Index Panel* (4) lists the most prevalent keywords (identified using frequency of keywords within the discussion) and most active contributors (number of posts) in descending order from left to right. The *Content-contributor Overview Bar* (12) provides the overview of where in the discussion a user talked about a particular keyword. The "*k*" label provides insight in the distribution of the presence of a given keyword in the discussion, selected from the *Keywords* in the *Index Panel*. Analogously, the "*u*" label does the same for a given user, selected from the *Users* in the *Index Panel*. The *Timeline* (13) gives the overview of when the posts were posted, thus helping to determine the distribution of

activity in the discussion over time. The *Posts Length* section (9) shows the amount of elaboration in terms of numbers of words in each of the posts, the line length being proportional to it. The *Replies Level* section (10) gives the overview of the users' depth of engagement – the length of the blue bar represents the depth of the reply.

Finding topics of interest: the key feature for enabling this capability is the *Keywords* in the *Index Panel* (4). By selecting a keyword from *Keywords*, only posts that contain that keyword are displayed in the *Discussion* section (3), keeping their sequential order. A second level of filtering topics of interest is introduced in the *Posts Keywords* (identified by tf-idf metric where a post is considered to be the document and the discussion the corpus) section (8), where the user can see which are most important keywords for that particular post.

Summary of opinions: the key feature in this category is the *Index Panel* (4) again. By selecting a keyword from *Keywords* and a discussion participant from *Participants*, the user can filter the posts that mention that keyword and come from that participant.
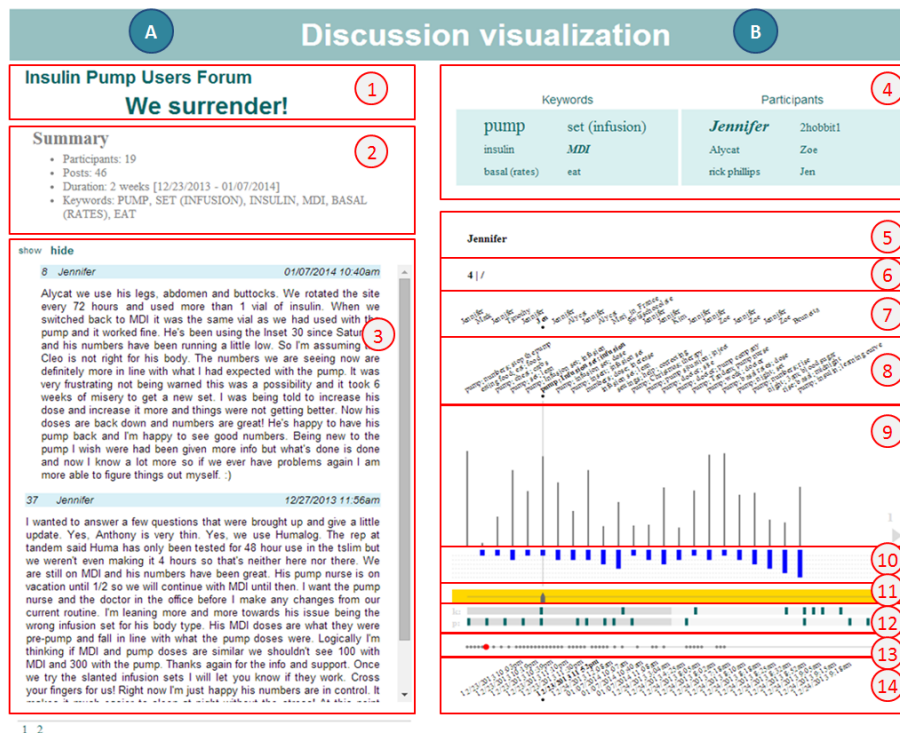

Figure 1. A snapshot of the DisVis user interface in use with functionality markers: (1) Category and Heading; (2) Summary; (3) Discussion; (4) Index Panel; (5) Interaction; (6) Context; (7) Authors; (8) Posts Keywords; (9) Posts Length; (10) Replies Level; (11) Slider; (12) Content-contributor Overview Bar; (13) Timeline; (14) Time and Date.

In addition to these three main groups of features, the remaining features of DisVis allow for integration between the two main segments (A and B). The most important of these is the *Slider* (11). Its cursor can be moved to point to a given post in its graphical presentation. Here, each post is represented by a vertical line segment, preserving the sequential order in the standard discussion representation (3); the different characteristics of these segments indicate their length (line height - 9), keywords (8), author (7), chronological position (13), time and date (14), and depth of engagement (rectangle height -10). To provide the social and informational context of the selected post, the *Interaction* feature (5) shows the set of people the author of that post had interacted the most in the discussion and through *Context* (6) it shows the posts that replied to it and the post it replied to.

Usage scenario. Elizabeth was recently diagnosed with Type 1 diabetes and was recommended an insulin pump as a treatment. She has heard about TuDiabetes and decides to visit the forum to learn more about it. She finds a discussion thread (displayed in the traditional way) that appears to be focused on insulin pumps, but finds that it has 5 pages of posts (a minimum of 41 posts). Elizabeth is not sure whether this is the right thread to focus on and decides to use DisVis to learn more about who participated and what was discussed. From the *Summary* section (2), Elizabeth can immediately see that 19 members of TuDiabetes participated in the discussion contributing 46 posts over only 2 weeks, and that the main topics included *insulin pump, infusion sets, insulin, MDI, basal rates* and

eating. She notices that one of the keywords included in the *Index Panel* (4) refers to the topic she is interested in, "pumps", and decides to investigate further. Elizabeth also notices that many of the posts have multiple replies by looking at the *Replies Level* section (10), which suggests to her that participants engaged in a discussion over the topics of mutual interest. This high-level overview gives Elizabeth a good sense of the complexity, richness and duration of the discussion, as well as participants' engagement. The initial post of the discussion is an appeal from Jennifer for troubleshooting unsuccessful insulin pump treatment and finding alternatives. After reading that, Elizabeth wants to explore Jennifer's opinions about pumps and whether they changed during the discussion. In the *Index Panel* (4), she selects "Jenifer" as the user, and selects the keyword "pump" from the *Index Panel* (4) as her topic of interest. In the *Discussion* section (3), she sees filtered posts that contain "pump" as a keyword and Jennifer as the author. She then reads those posts and formulates the opinion that Jennifer had on pumps. She does that for 2hoobit1 and other participants. However, she also notices that participants are mentioning MDI as an alternative treatment. Elizabeth goes through an analogous process to find out what they think about MDIs. As a result of these explorations, Elizabeth forms a pretty good idea of the different perspectives on pumps that were expressed in this discussion, and how members' perspectives changed over time. She also identifies an alternative treatment – MDI, that she can read more about in the future. She makes a mental note of the names of active participants and decides to pay attention to their posts in the future. And she is able to accomplish all of that in a matter of minutes, without reading every single post in this discussion thread.

Study Design

Sample and setting. The participants of the study were recruited among members of the TuDiabetes OHC using announcements on its home page and with the help of the community leadership. Being a member of this OHC was the only eligibility criterion for the study. The study was conducted on-line, using the Join.me shared screen platform for interacting with a discussion using the tool or within its original environment in TuDiabetes.

Assessment measures. There were three types (subsets) of questions used to evaluate the tool: S1, S2 and S3, with a total of 15 questions: a) *general overview of the discussion (S1)*, related to providing a summary of the discussion - finding the key concepts discussed and most active participants, activity of the discussion, etc. (e.g. What were some key concepts discussed here?); b) *finding topics of interest (S2)*, related to the ability to filter posts about a given topic (e.g. Who contributed the most posts that talked about infusion sets?); c) *opinion synthesis (S3)*, related to finding what different opinions are on a given topic (e.g. How would you summarize the general opinion about MDI, do you think most participants think it's a good idea, bad idea, or is there a split in opinions?). To assess participants' *accuracy*, we developed a *gold standard answers* (GS) for each of the 15 questions. For purely objective questions (e.g.: How many individuals participated in the discussion?) the answers were generated by the members of the research team. For questions that required domain knowledge and synthesis of opinions (e.g.: What were the key concepts discussed in this post?), the gold standard answers were generated by a clinician (a doctoral student in nursing with experience in diabetes self-management). We took the following approach to measure the accuracy based on the developed gold standard. For answers that had a single discrete value, the score could be either 0 or 1 (accurate=1, inaccurate=0) and for those that had several discrete values, we calculated the Jaccard[*] similarity coefficient between the values provided in the answer and the GS answer. For those that were qualitative and more descriptive, we assessed the similarity by Jaccard to the GS based on the key points covered in the answer compared to the GS (each key point being a discrete value, hence the user's answer and the GS answer being two sets of discrete values). To assess participants' *time to answer questions*, the researchers reviewed audio records of the sessions and measured time (in seconds) from the end of the question to the end of the final answer provided by the participant. The participants were given a maximum of 5 min. to start answering a question. If in that time frame an answer was not provided, their accuracy score would get set to 0 and time to answer would get set to 300 sec (this never happened however).

Study procedures. Upon enrollment, the participants were randomized into the experimental condition or the control condition (*N*=5 for each group). The participants in the control group were presented with a discussion thread within its original environment, TuDiabetes, and were asked a set of questions about the information contained in the discussion. After they answered all the questions, they were presented with DisVis, received a short tutorial, and asked for their initial feedback about the tool. The participants in the intervention group were presented with the same discussion thread within the DisVis interface. After receiving the same tutorial for the tool, they were asked the same questions as the control group. Then, they were asked for their feedback on the design of the tool and their

---

[*] Jaccard similarity is a metric for measuring the overlap between two sets of discrete elements which can take a value between 0 and 1. A higher value means a bigger overlap.

experience using it. All evaluation sessions were done through the Join.me platform and audio recorded and transcribed verbatim for analysis.

Analysis. To find out whether DisVis had an effect, we performed a two-tailed two sample t-test for each of the two outcomes: time and accuracy. Additionally, we wanted to see the tool's effect for each type of questions separately: S1, S2 and S3. Due to the small number of questions in each of the question categories, we used non-parametric Mann-Whitney U Test (Wilcoxon Rank Sum Test) to determine the statistical significance of our findings. For the ranking, we used the average performance of each participant for the given set of questions. For the qualitative data collected in the study (participants reflections on the tool and their experience with it) we used inductive thematic analysis. The recorded evaluation sessions were transcribed and open coding approach was taken to help us identify and categorize the prevalent themes from the user feedback. The first author (DN) coded all transcripts independently. The emerging categories were then discussed during meetings between the authors and iteratively refined in a collaborative process.

Results

Ten participants were recruited to participate in the study ($N=5$ in each group), 6 male and 4 female. More than half of the participants were above 60 years of age; diabetes diagnoses included both Type 1 and Type 2 diabetes having experience living with the disease ranging from 3 to 45 years (half of them for more than 20 years). The experience with the TuDiabetes.org ranged from only several months to being present in the community since its inception – almost 8 years. However, the vast majority of the participants were regular visitors of the forum, both for reading and contributing purposes; only one participant self-described as a sporadic user of TuDiabetes.

Quantitative

Table 1. Average time to answer and average score of answers for a set of questions and all questions overall.

category	measurement	group	mean	sd	significance test ($\alpha=0.05$)
S1	Time (**sec**)	Control	**70.16**	42.23	**(U=4, p=0.095)**
		Intervention	**34.73**	23.37	
	Accuracy	Control	**0.34**	0.06	**(U=1, p=0.016)**
		Intervention	**0.59**	0.14	
S2	Time (**sec**)	Control	**94.3**	41.72	**(U=4, p=0.095)**
		Intervention	**47.9**	43.30	
	Accuracy	Control	**0.6**	0.22	**(U=12.5, p=1)**
		Intervention	**0.6**	0.22	
S3	Time (**sec**)	Control	**40.7**	16.10	**(U=12, p=1)**
		Intervention	**49.8**	36.37	
	Accuracy	Control	**0.37**	0.19	**(U=1, 0.016)**
		Intervention	**0.77**	0.11	
All	Time (**sec**)	Control	**65.52**	46.77	**(t=1.80, p=0.082)**
		Intervention	**40.51**	26.38	
	Accuracy	Control	**0.38**	0.28	**(t=2.41, p=0.023)**
		Intervention	**0.64**	0.31	

Noticeable differences were observed in the advantage the tool brings over the control group. The tool was shown to highly improve the accuracy of and time to answering questions. Also, participants in the intervention group outperformed those in the control group for some, but not all of the question sets tested. From Table 1 (All) we can see that the average time to answer a question for the control group was 65.52 sec (*sd*=46.77 and median of 63.8 sec) as compared to 40.51 sec (*sd*=26.38 and median of 37.6 sec) for the intervention group. The average accuracy score for the control group was 0.38 (*sd*=0.28 with median of 0.35), and 0.64 for the intervention group (*sd*=0.31 with median of 0.8). The t-test showed that there is an effect for the accuracy of answering when using the tool: the accuracy when using the tool being 68% higher than without the tool (p-value = 0.023). However, the positive trends in reducing the time to answer questions by 38% was shown to be statistically insignificant (p-value = 0.082). The tool was shown to perform better than or at least as good as the control in all cases, except for time to answer questions from S_3. However, statistical significance was observed in only two of the cases in respect to question types: improving the accuracy of questions of type 1 (S_1) and type 3 (S_3), both with *p-value = 0.016*.

Qualitative

The focus of the qualitative interviews was on the participants' subjective impressions of the tool and their experience using it. The analysis of the transcripts suggested that the participants' experiences with the tool, and

with the forum at large, differed depending on their overall goal. In some cases, the participants came to the forum without any specific questions, but rather to look around for what's new, which included browsing of recent additions to the discussion threads, and skimming of the new content. However, in other cases, the participants came to the forum with a particular question in mind and their interaction focused on identifying relevant information through reading relevant posts and threads, and comparing different opinions on the topic. Notably, these two scenarios were interchangeably utilized across most of the participants. Also, and not surprisingly, these two usage scenarios were associated with very different expectations for how to optimize individuals' experiences and the desired set of functions within the forum. Below, we describe the participants overall assessment of the DisVis tool in terms of usefulness and usability, as well as the more specific comments from the perspective of browsing and searching for specific information.

Overall assessment of the DisVis tool. Overall, the majority of participants agreed with the general premise of the tool, and found its goal of increasing readability of discussion threads and promoting individuals' ability to explore threads to be promising and potentially useful. P1: "*Well, I think it's a great idea because I've had some struggles reading discussions, to be honest…*" P8: "*Well, for sure the tool is good to have since it gives you a clear idea about what's going on.*" However, together with these positive general perceptions, the participants had many critical comments regarding the tools' current design. Many thought it was too complicated and overthought and provided too much information, at times, without a clear purpose. P5: "*And I back up a little bit more, from my way that I use the site, this stuff would be more noise than signal.*" Notably, those who could not find personal use of the tool proposed that it might be most useful for those who started the thread, as well as for moderators, analysts, and researchers. P9: "*I mean, if you are a scientist analyzing it [the discussion] and you are trying like to pinpoint things like you've been asking me, than I can understand it [the reason for using the tool], but just for an average reader – I don' think I would use it.*"

Finally, the participants were split in their opinions about the tool's intuitiveness. Most users were able to successfully navigate the tool after the brief tutorial. However, several noted that they would require more experience using the tool to become proficient, which indicated limitations in its intuitiveness. P1: "*… there are also I think some struggles at the beginning, but now I think with a bit like getting used to it, it's not that difficult to use.*" P2: "*Well, you know I understood the instructions when you gave them, but putting them into practice, I think after I practice with it a couple or three times, I would have a down path, but just based on one run through, I don't -- I would need more time to learn it and practice with it.*" To address this problem specifically, given the number of different features available to users, several participants suggested the use of tool tips to remind them of the purpose of different features.

Exploratory browsing. While talking about their experience using the tool, many participants contextualized them within their overall experiences with TuDiabetes and its particular platform[*]. As we mentioned above, these experiences fell within two distinct categories: *exploratory browsing* and *focused information search*. We found that many of the specific comments in regards to DisVis's usefulness were also made in reference to its ability to support either browsing, or focused search.

Most participants made positive comments in regards to the tool's ability to support browsing. For example, the participants found the overview of the discussion, *Summary* statistics of the discussion and the *Timeline,* to be particularly helpful. P2: "*I like the dates, I like the date of the -- I really like knowing the timeframe of the thread and the number of participants. All of that summary information is very good.*" On the other hand, the *Slider* did not live up to the expectations. For example, the interactive graphical part of the visualization that involved the *Slider*, developed for navigational purposes for quick browsing, caused the most negative reactions. This was primarily due to the density of information, small font size used in that area and poor adjustability for different screen sizes, especially for small screens. The participants had split opinions about the list of participants and their number. Some of the participants found those useful for formulating opinion summaries or knowing what positions to expect in the discussion; others found them somewhat irrelevant.

More generally, however, some participants questioned the tool's attention on a single discussion thread as the only focus in the visualization. When describing their browsing experiences, these participants often talked about the need to see a discussion within the context of the forum, rather than as an isolated instance. P3: "*I'm more inclined to look at all the discussions, all the responses.*" These participants were likely to perceive individual threads as existing in a symbiotic relationship with other discussions that are on the same topic or have a certain amount of

---

[*] At the time of the study, TuDiabetes used Ning as its platform; since then the forum switched to Discourse

overlap in the topics they cover, alluding for a broader relevant context when reading a particular discussion. This colored their perceptions of the tool's focus on a single discussion thread: P7: *"…so if you clicked on pump you would see a list here of all of the discussions' summaries of whatever, of, of whatever is related in, in another discussion, that, that would be good."* These participants wished to see the overlap between different discussions on the same topic, and to compare the topics between discussion threads. P10: *"Perhaps instead of, instead of this, maybe, on keywords, if you clicked on insulin or mdi or pump instead of showing where it shows up, here, what other discussions are mentioning pump, insulin and mdi…"* Participants were also interested to see *forum wide statistics*, a more quantitative description of what was going on in the forum: P3: *"I might want to compare this to discussions and see what the keywords are for both and see if I could cross reference some keywords. I might also want to know what the community's trending keywords are."* Finally, the participants proposed the following new features for more efficient skimming of discussions: highlighting keywords of interest in the text, finding the latest posts in the discussion, and labeling the read posts.

Focused information search. According to the participants, the tool was helpful in finding specific information of interest. The most-liked feature for this task was the Index Panel with its Keywords section. P2: *"I would love to sort of sort out especially with long threads the ones that really went back to the key points of the initial discussion, like in this case pump and insulin I would probably because I'm an insulin user. I would probably pick on insulin and pump. So it does help me …"* However, some participants thought they could not find much use of the Index Panel in real life, although it helped them answer the questions in the study. The participants also proposed specific enhancements to existing features. These included: more prominent highlighting of the selections in the Index Panel, and selection of multiple keywords for the purpose of building more complex Boolean queries.

Further, the participants made many comments regarding desired new features that could improve searching for specific information. These included features for identifying posts of relevance within a particular discussion (as was the case in DisVis) as well as for finding information of interest within the forum at large, or as some participants referred to it "starting from scratch". P10: *"So let's say I go in, onto the home page, and I have a question about the pump… you know… how do I find it from scratch?! Instead of being right here on this open page that's already started and everything, how would I get to that page from scratch?? – Do I see five different discussions going on, 20 different discussions and then, what would I do??"* For the first case, the participants proposed keywords (terms) that distinguish a discussion from other similar discussions, as well as a keyword cloud. P4:" *…I'm in the insulin pump users' forum so I would already expect there to be discussions about the pump and oftentimes, a lot of the discussions are going to include comments about the infusion set, about the insulin, about the basal rates and multiple daily injections. So I guess the question for me is how could we better define keywords that actually provide more relevance… […] Yeah, I think more about differentiating the discussions [in terms of keywords] because I think a lot of people will come out in these forums and will search for somebody having the same issue or same challenge that they have today."* For the second case, the participants strongly wished for a better search engine, claiming that the current one is simply not delivering the most relevant and recent discussions, and it seems to be more oriented towards individual posts within a discussion, rather than a discussion as a whole. P1:*"If you search for a key word it doesn't really show you what you are looking for, like searches for specific answers and I want the overall discussions , so that's not much of help…"* In addition, an ability to identify clusters of discussions was highly desired as a feature that can improve the searching process. Finally, the participants liked the idea of exploring different opinions about a topic of interest. However, there was no clear agreement regarding the best way to achieve this. Some wished for a condensed, potentially computer-generated summary; others felt that only reading multiple posts from the same person on a topic of interest could help to synthesize their perspective, an ability which comes through long time presence in the community and reading of the forum.

Discussion

Previously proposed tools for visualizing participant-generated content explored several different perspectives. Many tools focused on providing an overview of the entire volume of the content collected (e-mails or news group postings), rather than focusing on a specific conversation. These included visualizations of interpersonal interactions and topics covered in those interactions[22]; engagement of participants in discussions[23] and individuals' activity and volume of contributions over time[23]. Other tools took an approach more consistent with the one explored in this work and focused on providing overviews of individual discussions as well as more detailed information about them. Some of these tools placed greater emphasis on finding relevant posts based on topics of interest and participants, while also incorporating information regarding the sentiment of the topic[20]. Others were concerned with providing visual, often scatterplot-based depictions of clustered discussion content. These tools distinguished discussions and threads of conversation within the discussion, highlighted different topics and participants, and

importance of individual posts[21,25]. The evaluation studies of these tools focused on users' ability to find topics of interest, summarize, compare and contrast opinions, understand the sentiment about a topic, activity of the participants, their expertise and their interpersonal interactions and relationships. The authors' activity oriented studies showed that knowing authors behavioral patterns in terms of activity and types of contributions was highly valued for understanding their roles and domain knowledge in the community[22,23]. On the other hand, the discussion visualization studies found that users particularly liked the ability to find posts that match their topic of interest, which was found useful in answering questions related to summarizing opinions[20]. Here, however, the findings about the importance of authors' details were dependent on the task study participants were asked to complete[20,21].

Consistently with these previous studies, the participants in our study were positive about the general premise of an interactive tool for visualizing salient features of a discussion thread and exploring its different characteristics and attributes. We found that the participants viewed the tool in the context of two typical scenarios of use, corresponding to the two general types of engagement they experienced with the forum, a distinction not considered in the previous research. On one hand, they looked for features that could help to summarize the thread in a bird's eye view, highlighting such characteristics as most prevalent keywords and most active participants. On the other hand, they particularly liked the possibility to find posts of interest using the *Index Panel*, which helped in finding specific pieces of information and formulating opinions about a topic. This is consistent with other studies which found that users particularly liked the ability to find posts that match their topic of interest[20,21]. In addition, the participants valued an ability to view a temporal visualization of the discussion in the interactive timeline, an ability that has not received much attention in prior research, but was considered particularly useful in our study. These positive impressions were also supported by the results of the quantitative assessment: participants in the intervention group who used DisVis interface were able to answer questions related to the discussion thread with higher accuracy and in less time (question types S1 and S3). However, together with these positive impressions, the participants made a number of critical comments, particularly in regards to the tools' intuitiveness and clarity. Most importantly, they found the information display to be too dense and hard to read, and some of the interactive features, for example the Slider, too difficult to operate. This can be related to a finding from another study, where participants underperformed when exposed to both visualization of the discussion and text analytics compared to a standard discussion representation[21]. However, it contradicts our case in which, based on the quantitative findings, the complexity of the tool actually improved performance as compared to the baseline. From the perspective of general browsing, the participants questioned the tool's focus on a single discussion thread, and desired an overview of topics of interest throughout all discussion threads in the forum. In contrast to our work, this kind of an approach was predominant in VisOHC which focused on OHC moderatos[34]. From the perspective of focused search for information, participants found one keyword selection to be limiting and wished for an ability to construct more complex queries. In addition, they wished for an ability to see keywords selected within the Index Panel highlighted in the discussion's text.

The study helped us to formulate a number of design implications for the next generation of tools for visualizing discussions in OHCs. First of all, the study once again stressed the importance of designing tools that could be used by a variety of individuals regardless of their background, age, specifics of their disease, and computer literacy. Since these tools are likely to be used in an unsupervised fashion and without any possibility for training or instruction, they need to be intuitive and easy to learn. The study also generated a number of specific recommendations for features that could promote the two complementary ways of engaging with information. First, regarding the *browsing* experience, the desired features included: 1) discussion summary (statistics); 2) larger context, including immediate access to relevant discussions; 3) capabilities to estimate and locate the overlap between the relevant discussions; and 4) keeping track of what was read and what is new. Second, regarding the *focused information searching*, the desired features included: 1) description of the discussion that distinguishes it from the others 2) building more complex queries to search a particular discussion and 3) computationally assisted ways to summarize different opinions within the discussion.

In addition, the study generated new design directions for helping individuals engage with user-generated content within OHCs. These include: a) *identifying and highlighting relationships between different discussion threads, for example those that include similar topics* and b) *developing customizable mechanisms to summarize varying opinions on topics of interest*. However, in regards to summarizing opinions, the participants lacked clear consensus on whether they were comfortable relying on computer-generated summaries or whether they preferred tools that helped humans to synthesize opinions in a more efficient way. This division is reminiscent of the long-standing argument about the comparative benefits of computational agents and information visualization displays with direct manipulation[36]. Just as this debate came to the conclusion that an ultimate solution requires a combination of both

approaches, we suspect that future tools for supporting online discussion forums should incorporate both computational methods and information visualization techniques.

Study limitations. This study has a number of limitations. First, because the study relied on a real discussion thread captured within TuDiabetes, and because we did not screen participants on their previous exposure to this particular thread, it is possible that some of them saw it before the study. In fact, one of the participants mentioned that the thread looked familiar, but could not recall any specific details. In addition, because the study used only a single discussion thread, it is possible that the findings of the study are specific to this thread and may not generalize to other threads within the same forum or to other forums. In addition, the questions used for assessing the tool's efficacy may not be relevant to all individuals who frequent OHCs. In fact, few of these questions were found less than relevant by some of the study participants. Additionally, sharing the screen occasionally caused some difficulties with interaction on the participants' side for both study groups, which was more pronounced for those that needed to interact with the tool.

Conclusion

In this paper, we introduced DisVis – a tool for visualizing discussion threads in an OHC and discussed the results of a study examining DisVis's impact on individuals' ability to find information of interest and better understand discussion threads. The study showed that DisVis had a positive overall effect on individuals' ability to provide overview of the discussion, find topics of interest and summarize opinions. The tool increased the accuracy of the answers by 68% and reduced the time to answer questions by 38% (the latter being statistically not significant). The study also showed that the community members are open to these types of solutions and suggested a set of enhancements and new features for the tool that we plan to address in our future work.

References

1. Grace J. Johnson PJA. Neo-tribes: the power and potential of online communities in health care. Commun ACM. 2006;49:107–13.
2. Griffiths F, Cave J, Boardman F, Ren J, Pawlikowska T, Ball R, et al. Social networks--the future for health care delivery. Soc Sci Med 1982. 2012 Dec;75(12):2233–41.
3. Denecke K, Nejdl W. How Valuable is Medical Social Media Data? Content Analysis of the Medical Web. Inf Sci. 2009 May;179(12):1870–80.
4. Fox S, Purcell K. Social Media and Health, Pew Research Center. June, 2009.
5. Fox S, Purcell K. Social Media and Health, Pew Research Center. March, 2010.
6. Chen AT. Exploring online support spaces: Using cluster analysis to examine breast cancer, diabetes and fibromyalgia support groups. Patient Educ Couns. 2012 May;87(2):250–7.
7. Antheunis ML, Tates K, Nieboer TE. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ Couns. 2013 Sep;92(3):426–31.
8. Mamykina L, Nakikj D, Elhadad N. Collective Sensemaking in Online Health Forums. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2015. p. 3217–26. (CHI '15).
9. Vlahovic TA, Wang Y-C, Kraut RE, Levine JM. Support Matching and Satisfaction in an Online Breast Cancer Support Community. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2014. p. 1625–34. (CHI '14).
10. Zhang Y. Contextualizing Consumer Health Information Searching: An Analysis of Questions in a Social Q&A Community. In: Proceedings of the 1st ACM International Health Informatics Symposium [Internet]. New York, NY, USA: ACM; 2010. p. 210–9. (IHI '10).
11. De Choudhury M, Morris MR, White RW. Seeking and Sharing Health Information Online: Comparing Search Engines and Social Media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2014. p. 1365–76. (CHI '14).
12. Annie T Chen LS. Topic Modeling and Network Visualization to Explore Patient Experiences. 2013;
13. Chen A. Patient Experience in Online Support Forums: Modeling Interpersonal Interactions and Medication Use. 51st Annu Meet Assoc Comput Linguist Proc Stud Res Workshop. 2013;16–22.
14. Civan A, McDonald DW, Unruh KT, Pratt W. Locating Patient Expertise in Everyday Life. In: Proceedings of the ACM 2009 International Conference on Supporting Group Work. New York, NY, USA: ACM; 2009. p. 291–300. (GROUP '09).

15. Civan-Hartzler A, McDonald DW, Powell C, Skeels MM, Mukai M, Pratt W. Bringing the Field into Focus: User-centered Design of a Patient Expertise Locator. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2010. p. 1675–84. (CHI '10).

16. Mukherjee S, Weikum G, Danescu-Niculescu-Mizil C. People on Drugs: Credibility of User Statements in Health Communities. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: ACM; 2014. p. 65–74. (KDD '14).

17. Fernandez-Luque L, Karlsen R, Bonander J. Review of extracting information from the Social Web for health personalization. J Med Internet Res. 2011;13(1):e15.

18. Liu Y, Xu S, Yoon H-J, Tourassi G. Extracting patient demographics and personal medical information from online health forums. AMIA Annu Symp Proc AMIA Symp AMIA Symp. 2014;2014:1825–34.

19. Nambisan P. Information seeking and social support in online health communities: impact on patients' perceived empathy. J Am Med Inform Assoc JAMIA. 2011;18(3):298–304.

20. Hoque E, Carenini G. ConVis: A Visual Text Analytic System for Exploring Blog Conversations. Comput Graph Forum. 2014 Jun 1;33(3):221–30.

21. Dave K, Wattenberg M, Muller M. Flash Forums and forumReader: Navigating a New Kind of Large-scale Online Discussion. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work. New York, NY, USA: ACM; 2004. p. 232–41. (CSCW '04).

22. Viégas FB, Golder S, Donath J. Visualizing Email Content: Portraying Relationships from Conversational Histories. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2006. p. 979–88. (CHI '06).

23. Viegas FB, Smith M. Newsgroup Crowds and AuthorLines: visualizing the activity of individuals in conversational cyberspaces. In: Proceedings of the 37th Annual Hawaii International Conference on System Sciences, 2004. 2004. p. 10 pp.–.

24. Pascual-Cid V, Kaltenbrunner A. Exploring Asynchronous Online Discussions through Hierarchical Visualisation. In: Information Visualisation, 2009 13th International Conference. 2009. p. 191–6.

25. Narayan S, Cheshire C. Not Too Long to Read: The Tldr Interface for Exploring and Navigating Large-Scale Discussion Spaces. In: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences. Washington, DC, USA: IEEE Computer Society; 2010. p. 1–10. (HICSS '10).

26. Smith MA, Fiore AT. Visualization Components for Persistent Conversations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2001. p. 136–43. (CHI '01).

27. Cui W, Liu S, Tan L, Shi C, Song Y, Gao Z, et al. TextFlow: Towards Better Understanding of Evolving Topics in Text. IEEE Trans Vis Comput Graph. 2011 Dec;17(12):2412–21.

28. Heer J, Boyd D. Vizster: visualizing online social networks. In: IEEE Symposium on Information Visualization, 2005 INFOVIS 2005. 2005. p. 32–9.

29. Welser HT, Gleave E, Fisher D, Smith M. Visualizing the signatures of social roles in online discussion groups. J Soc Struct. 2007;8(2):564–86.

30. Liu B, Hu M, Cheng J. Opinion Observer: Analyzing and Comparing Opinions on the Web. In: Proceedings of the 14th International Conference on World Wide Web. New York, NY, USA: ACM; 2005. p. 342–51. (WWW '05).

31. Moghaddam RZ, Bailey BP, Poon C. Ideatracker: An Interactive Visualization Supporting Collaboration and Consensus Building in Online Interface Design Discussions. In: Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part I. Berlin, Heidelberg: Springer-Verlag; 2011. p. 259–76. (INTERACT'11).

32. Chen AT. Information and Emotion in Online Health-Related Discussions: Visualizing Connections and Causal Chains. Medicine 2.0. 2012.

33. Chen AT, Zhu S-H, Conway M. What Online Communities Can Tell Us About Electronic Cigarettes and Hookah Use: A Study Using Text Mining and Visualization Techniques. J Med Internet Res. 2015;17(9):e220.

34. Kwon BC, Kim S-H, Lee S, Choo J, Huh J, Yi JS. VisOHC: Designing Visual Analytics for Online Health Communities. IEEE Trans Vis Comput Graph. 2016 Jan;22(1):71–80.

35. Hutchinson H, Mackay W, Westerlund B, Bederson BB, Druin A, Plaisant C, et al. Technology Probes: Inspiring Design for and with Families. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2003. p. 17–24. (CHI '03).

36. Shneiderman B, Maes P. Direct Manipulation vs. Interface Agents. Interactions. 1997 Nov;4(6):42–61.