Received: 23 April 2020 | Revised: 12 October 2020 | Accepted: 15 November 2020

DOI: 10.1002/asi.24436

RESEARCH ARTICLE

JASIST WILEY

Predicting users′ continued engagement in online health communities from the quantity and quality of received support

Xiangyu Wang[1] | **Andrew High**[2] | **Xi Wang**[3] | **Kang Zhao**[4]

[1]Interdisciplinary Graduate Program in Informatics, The University of Iowa, Iowa City, Iowa

[2]Department of Communication Arts and Sciences, Pennsylvania State University, University Park, Pennsylvania

[3]School of Information, Central University of Finance and Economics, Beijing, China

[4]Tipple College of Business, The University of Iowa, Iowa City, Iowa

Correspondence
Kang Zhao, The University of Iowa, Iowa City, IA 52242-1994.
Email: kang-zhao@uiowa.edu

Abstract

Online health communities (OHCs) have been major resources for people with similar health concerns to interact with each other. They offer easily accessible platforms for users to seek, receive, and provide supports by posting. Taking the advantage of text mining and machine learning techniques, we identified social support type(s) in each post and a new user's support needs in an OHC. We examined a user's first-time support-seeking experience by measuring both quantity and quality of received support. Our results revealed that the amount and match of received support are positive and significant predictors of new users' continued engagement. Our outcomes can provide insight for designing and managing a sustainable OHC by retaining users.

1 | INTRODUCTION

The ubiquity of the Internet and the rising usage of mobile devices have changed the way people access information. More and more people seek health-related information online (Fox, 2014) and interact with others about health issues in technologically mediated channels (Chou, Hunt, Beckjord, Moser, & Hesse, 2009). According to the Pew Research Center, 80% of all Internet users look for health-related information online, 72% of adults in the United States use the Internet to search specific diseases and treatments, and 26% of adult Internet users read other's health experiences. Health information from the Internet could come from different types of platforms, including general-purpose or specialized information repositories, including Wikipedia.org and WebMD.com, as well as community-based platforms such as online health communities (OHCs), where users network with peers with similar health problems or concerns (Xu, Zhou, Zhang, & Hendler, 2018).

The pillar of OHCs is the exchange of social support, which includes but also goes beyond health information (Wang, Zuo, & Zhao, 2015). Like other online communities, OHCs often offer anonymity and privacy for users (Hwang et al., 2010), so that even strangers can seek and provide support without fear or stigma (S. E. Caplan & Turner, 2007). The Internet is not limited by geographical and time constraints (Barak, Boniel-Nissim, & Suler, 2008), making it easier for users of OHCs to access support than people who utilize face-to-face support groups (Katz, Rice, & Aspden, 2001), especially for those who suffer from rare, chronic or debilitating conditions (Hawn, 2009). In addition, (Walther & Boyd, 2002) reasoned that online social support allows enhanced control over how an interaction unfolds, and increases the social proximity between support seekers and potential providers. Research has also revealed that OHCs are helpful for users to learn symptoms or treatments and connect with peers to better understand their own health conditions (Wicks et al., 2010). The benefits of social support have been well documented in the literature. For example, positive social interactions between support seekers and providers in OHCs can be helpful and therapeutic (De Choudhury & De, 2014). Social support is associated with improved psychological well-being (Chou et al., 2009; Kim et al., 2012; Namkoong et al., 2010; Qiu et al., 2011; Yoo et al., 2014) and better health outcomes (Eaker, 2005; Maunsell, Brisson, & Deschěnes, 1995).

The benefits of supportive communication from OHCs cannot be realized if support seekers do not receive what they need and stop their participation in the group. The amount of support is one intuitive measure of received support. Besides the amount of received support, the match of social support, or the extent to which provided support matches people's needs (Cutrona, Shaffer, Wesner, & Gardner, 2007), also matters. The theoretical foundation of matching social support can be traced back to the person-environment fit model (R. D. Caplan, 1987), which proposes that the effectiveness of an exchange of resources depends on the fit between recipients' needs and what they receive (Matire, Stephens, Druley, & Wojno, 2002). A potential mismatch occurs when actual needs of support seekers are different from providers' understanding or what seekers communicate to others (Arora, Finney Rutten, Gustafson, Moser, & Hawkins, 2007). A related approach to understanding the efficacy of supportive interactions focuses on support gaps, or the discrepancy between the amount of different types of support people desire or seek and what they receive in a particular interaction (High & Crowley, 2018; High & Steuber, 2014). A discrepancy between the support people desire and receive determines whether a gap exists, and the size of that gap shapes the outcomes of an interaction (Crowley & High, 2020). Support gaps exist between users and the responses they receive from online support groups, and support gaps often correspond with negative outcomes (Crowley & High, 2020; High & Steuber, 2014). For instance, empirical studies have found that mismatched social support can lead to poor physical and mental health conditions among those who need support (Reynolds & Perrin, 2004; Siewert, Antoniw, Kubiak, & Weber, 2011). Such negative outcomes may affect users' satisfaction with OHCs; therefore, it is important to better understand users' support needs and whether their needs are satisfied in an OHC (Arora et al., 2007).

Like many other online communities, OHCs also face user "churns"—users leaving the community. Because a successful and supportive OHC depends heavily on users' long-term engagement (Wang, Zhao, & Street, 2017), user retention is imperative for an OHC, especially when most OHCs do not offer any monetary incentive for users to stay. User churn often occurs at an early stage of a user's online participation (Graham et al., 2017; Wang, Zhao, Zhou, & Street, 2020), especially among those who fail to get relevant social support (Yang, Kraut, & Levine, 2017). Therefore, OHCs need to make sure new users are satisfied with their early experiences.

In this paper, we used data collected from a popular peer-to-peer OHC and investigated how users' first-time support-seeking experience in an OHC, combined with the quantity and match of the responses they receive, can predict their participation in this context. Mining large-scale data of an OHC for cancer survivors, we first adopted text mining to identify the types of social support activities in each post. This helped us identify a new user's social support needs and enabled us to measure the quantity and quality of social support she or he received from the OHC. We then documented that both the amount and match of social support received during a user's first-time support-seeking experience are positive predictors of users' subsequent participation in the OHC. The outcomes of this research have implications for an OHC to better manage a user's first-time experience and improve its user retention efforts. Specifically, understanding users' early-stage online participation and associated factors help to design a more effective OHC and encourage continuing participation in the OHC.

2 | RELATED WORK

2.1 | Social support in OHCs

Social support describes the exchange of resources between a provider and recipient (Shumaker & Brownell, 1984). People feel supported when they are loved, respected, cared for, or a member of a nurturing community (S. Cobb, 1976). The benefits users experience from supportive interactions are based on what and how they seek support (Barbee & Cunningham, 1995; Andrew C High & Crowley, 2018), as well as the support they receive. The amount of support received has been quantified as the number of comments to a thread (Yang et al., 2017).

Besides the amount of received support, the match of support is an important dimension of the quality of a supportive interaction. To more accurately measure the efficacy of received support, researchers need to consider the goals or purpose of posts that initiate threads in OHCs. People those who regard themselves to be less healthy, more distressed, or having experienced cancer are more likely to be involved in social support groups (Chou et al., 2009). In OHCs, people seek and provide several different types of support, thereby making the idea of matching people's desire for support is more important. Although (Cutrona & Suhr, 1992) acknowledged five distinct types of support, including informational, emotional, esteem, network, and tangible support, informational and emotional support are the most common types of support people encounter online (High, Jennings-Kelsall, Solomon, & Marshall, 2015; Rains, Peterson, & Wright, 2015). (Goldsmith, 1994) also argued that people in distress most need informational or emotional support. (Bambina, 2007) defined four types of social support in health-related communication: emotional support, informational support, companionship, and instrumental support. Emotional support involves sharing expressions of love, sympathy, encouragement, affection and understanding. Informational support is related to the exchange of advice, information and knowledge for related needs. Companionship is about informal chattings, discussion of daily life or other social activities that are not directly related to health. Instrumental support refers to tangible support, such as transporting others to hospitals or cooking. Although emotional, informational, and instrumental support provide resources to users directly, companionship support makes individuals feel valuable as they become a part of a group. Cancer patients, in particular, value informational support (Linden & Vodermaier, 2012), and emotional support is relevant to most stressors people face (MacGeorge, Feng, & Burleson, 2011). Receiving informational and emotional support sometimes also have different trajectories for people experiencing distress. (Jacobson, 1986) points out the timing of social support also effects the effectiveness. First-time users' posts are typically support-seeking in nature, new users often expect to receive informational support, whereas existing members long for emotional support (Yang et al., 2017). Therefore, it is valuable to examine how informational and emotional support might differently shape people's continued participation in OHCs. Prior research has also documented the value of expressing companionship in OHCs, even calling it the key to sustaining these communities (Wang et al., 2017). Among patients coping with dialysis, receiving companionship corresponded with reduced depression, and a gap between the amount of companionship people desired and received was associated with both increased depression and mortality risk (Thong, Kaptein, Krediet, Boeschoten, & Dekker, 2007). Because of their importance, we focus on informational and emotional support along with companionship in the current study.

In the context of social support in OHCs, the optimal match and support gaps frameworks indicate a clear alignment between the recipient's needs and the provider's response to those needs (Shumaker & Brownell, 1984). In contrast, a support gap or mismatch of social support is characterized by a discrepancy between the types of support desired and received. Although conceptually intuitive, it can be challenging to measure the match of social support. Most empirical studies used user satisfaction collected by surveys or interviews as a proxy to evaluate a social support match (Reynolds & Perrin, 2004; Vlahovic, Wang, Kraut, & Levine, 2014). Such approaches may suffer from users' inaccurate recall of memories, social desirability biases, and small sample sizes. In addition, some of these studies involve coding messages for the types of support that are sought or provided, and coding is often a time and work-intensive procedure, particularly when operating at the scale of an OHC. These limitations can be addressed by mining user-generated content, which is often available in large scales in OHCs.

Many studies have examined unstructured user-generated content using computational methods (e.g., machine learning methods). Information extraction tools (e.g., based on generic dictionaries) have also been used to differentiate seeking and providing support without training machine learning classifiers. However, the performance of such generic tools in different contexts may vary and was rarely evaluated (Lv, Ng, Lee, Sun, & Yeung, 2008). Thus, computational methods have been deployed to identify social support categories from user-generated content by extracting lexical features, sentiment features, and topic features (De Choudhury & De, 2014; Xi Wang et al., 2017; Y.-C. Wang, Kraut, & Levine, 2012; M. Zhang & Yang, 2017). Recent studies use word embeddings (e.g., Word2Vec) (Mikolov, Chen, Corrado, & Dean, 2013) to capture both semantic and syntactic features from user-generated content in OHCs (Khanpour, Caragea, & Biyani, 2018; S. Zhang, Grave, Sklar, & Elhadad, 2017). Such automated classification of types of social support offers new opportunities to measure support matching at a more fine-grained level.

2.2 | Users' participation in OHCs

Previous studies on predicting user participation in OHCs relied on two types of data. The first type aggregated the level of users' online activities into measures such as the number of threads initiated and the number of replies posted (Jones et al., 2011). Some studies also built various social networks among users and used individual centralities (N. K. Cobb, Graham, & Abrams, 2010; Healey, Hoek, & Edwards, 2014; Sudau et al., 2014; Zhao et al., 2016), or their social network dynamics (Graham et al., 2017), to predict user participation in OHCs. The second type examined user-generated content with either manual coding (Chuang & Yang, 2014) or machine learning (Xi Wang, Zhao, & Street, 2014). For example, research has found that social support can predict users' continued participation in offline social support groups (Ganz et al., 2002; Helgeson, Cohen, Schulz, & Yasko, 2000) as well as in OHCs. Meanwhile, specific types of support mined from user-generated content enable more fine-grained analyses describing how different types of support contribute to user participation. For instance, users' own activities in seeking emotional support and getting involved in companionship are better predictors of long-term engagement in an OHC than informational support (Wang et al., 2017; Y.-C. Wang et al., 2012). A user's social support activities could also change when they are more engaged in an OHC (Chuang & Yang, 2014; Wang et al., 2015).

Despite the valuable findings they generated, existing studies on users' participation have three limitations that we attempt to address in this paper: *First*, existing studies assumed all threads were posted to seek support (Chancellor, Hu, & De Choudhury, 2018); however, it is conceivable that the purpose of initiating a post varies by individuals. In the current study, we examine users whose initial posts started with strong intentions to seek support in OHCs, rather than assuming all users who start a thread are support seekers. *Second*, previous studies on users' participation mainly focused on users' lifelong participation (Ma et al., 2017; Wang et al., 2017; Y.-C. Wang et al., 2012). Although long-term participation is valuable, less attention has been paid to the immediate effect of social support exchanges on user's early stage participation. Recent studies demonstrated that members' behavior in their first month in an OHC effect their continued engagement (Ma et al., 2017); therefore, understanding users' early experience in an OHC could be crucial to preventing or limiting user churn. To gain insight into the value of social support provided in users' early engagement, we examine the effects of their first-time seeking support on their subsequent participation in different periods. *Third*, as mentioned earlier, most previous studies measured social support match by proxy variables. In these studies, social support match is measured by a binary variable, which leads to a loss of information. To remedy this issue, we measure social support received from initial posts according to both their quality and quantity.

3 | METHODS

The dataset for this study consists of user-generated posts from Breastcancer.org, a popular peer-to-peer OHC for breast cancer survivors. It includes all public posts from this community between October 2002 and August 2013. There are around 2.7 million posts organized into 93,453 threads, contributed by around 50,000 users. Users of the OHC in breastcancer.org agree that information they provide may be "read, collected, and used by others who access them." They can also ask the site to remove any information they shared.[1]

3.1 | Social support detection

Social support can be of different types. We adopted the typology of social support proposed by (Bambina, 2007)—emotional support, informational support, companionship, and instrumental support. Empirical studies have shown that instrumental support is not common in OHCs because it is limited by geographical constraints (Wang et al., 2017), and it often occurs through private communication channels such as email, text messaging, and so forth, instead of a public forum. Hence, we excluded instrumental support from this study. With the three types of social support, we defined five types of social support activities: Seeking Emotional Support (SES), Seeking Informational Support (SIS), Providing Emotional Support (PES), Providing Informational Support (PIS), and Companionship (COM). We did not differentiate the seeking and providing of companionship because everyone involved in companionship activities is seeking and providing this type of support at the same time. Table 1 shows[2] example posts from each of the five social support categories.

The first step of our analysis was to classify post content into different social support categories. It is impractical for humans to annotate the 2.7 million posts in our data. Machine learning methods were applied to automatically detect social support categories based on the content of posts. We recruited 5 human annotators to label 1,333 randomly selected posts from this dataset. To ensure the quality of annotations, we added 10 posts annotated by domain experts to the pool of posts as quality-control posts. For each post, we only accepted results from annotators whose performance on the 10 quality-control posts was among the top 3. The results from the other 2 annotators were discarded. Then, majority votes among the top 3 annotators were used to determine whether a post was related to a category of social support. Note that a post can belong to more than one category.

Unstructured text from users' posts were represented by five groups of features. Following procedures developed by Wang et al. (Wang et al., 2017), we extracted basic features, lexical features, sentiment features and topic features from each post. To cover all texts in a post, we also added Word2Vec, a word embedding method, to represent each post's content with a vector (Mikolov et al., 2013). Table 2 lists all features for post classifications. We trained five classifiers, one for each of the social support activities. When training classifiers for SIS and SES, we oversampled instances from the minority classes with SMOTE (Chawla, Bowyer, Hall, & Kegelmeyer, 2002) because proportions of positive instances are low for the two types of social support. Then, we applied six different classification

TABLE 1 Excerpts of example posts for each social support category

Social support categories	Example posts
Seeking informational support (SIS)	*… do you follow any specific breast cancer blogs, … do you read a blog that you think we would enjoy reading?*
Seeking emotional support (SES)	*Cranky and depressed tonight … I will go back into my corner.*
Seeking informational support (SIS) and seeking emotional support (SES)	*… I was diagnosed with breast cancer in …. I've been feeling scared, anxious, very nervous… my doctor gave me Xanax, … what is going on with me? Anyone else feeling this way? … after chemo, I started Zoloft… I felt weird and a little down… will this pass? Any advice?*
Providing informational support (PIS)	*… the best hospital for cancer is Sloan-Kettering cancer Institute in new York, Johns Hopkins in Baltimore, Mayo Clinic in Rochester and MD Anderson in Houston… UCLA in Los Angeles is good. … the UCSF hospital in San Francisco is great …*
Providing emotional support (PES)	*… please find courage, strength and hope to guide you and yours through this grieving process. … my sincere condolences.*
Providing informational support (PIS) and providing emotional support (PES)	*Stay blessed … my suggestion is to wait till the doctor tells you what medicine you will take. … Herceptin, Xeloda and Tykerb do not cause hair loss… I hated to wear a wig but there are some cute ones … bless you if you … best of luck.*
Companionship (COM)	*Wishing you the best birthday ever … happy birthday!*

TABLE 2 Summary of features for post classifications

Groups	Features
Basic features	Whether the post is an initial post
	Whether the post is a reply by the author of the initial post
	Number of words in the post
Lexical features	Whether the post contains URLs
	Whether the post contains emojis
	Number of numeric numbers
	Number of pronouns (e.g., they, we, I)
	Whether the post contains negation word(s) (e.g., not, never, no)
	Whether the post contains name(s) of cities, states, or countries
	Whether the post contains names of drugs for breast cancer (from http://www.cancer.gov/cancertopics/druginfo/breastcancer)
	Whether the post contains breast cancer terminology (from http://www.breastcancer.org/dictionary)
	Whether the post contains verb related to advice (e.g., need, require, recommend, etc.)
	Whether the post contains emotional words (e.g., love, sorry, hope, worry, etc.)
	Whether the post contains words related to seeking behaviors (e.g., anybody, question, wonder, etc.)
	Whether the post contains words related to daily life (e.g., vacation, joke, run, walk, etc.)
Sentiment features	Frequency of words with positive and negative sentiment obtained from (Hu & Liu, 2004)
	Objectivity and subjectivity scores obtained from python library TextBlob
Topic features	Topic distributions derived from LDA ($k = 20$)
Textual features	Vector representations obtained from Word2Vec

Note: Most features were adapted from (Wang et al., 2017).

algorithms using 10-fold cross-validation and compared their accuracies and area under the ROC (AUC). For word embeddings, we tried different embedding lengths (100 and 300), different window sizes (3 and 5) and both CBOW and Skip-gram models. We found that Skip-gram Word2Vec with window size = 3 and embedding vector

TABLE 3 The performance of different algorithms for social support category classifications

Social support	Metrics	Naïve Bayes	Logistic regression	Random forest	Decision tree	AdaBoost	XGBoost
COM	Accuracy	0.742	0.792	0.843	0.765	0.835	**0.850**
	AUC	0.751	0.869	0.912	0.749	0.896	**0.924**
PES	Accuracy	0.586	0.866	0.876	0.804	0.852	**0.878**
	AUC	0.745	0.884	0.911	0.692	0.864	**0.917**
PIS	Accuracy	0.692	0.832	0.841	0.774	0.832	**0.853**
	AUC	0.771	0.914	0.916	0.760	0.911	**0.927**
SES	Accuracy	0.580	**0.977**	0.978	0.947	0.974	0.978
	AUC	0.646	**0.750**	0.690	0.530	0.552	0.693
SIS	Accuracy	0.508	**0.906**	0.894	0.828	0.902	0.902
	AUC	0.699	**0.893**	0.807	0.597	0.834	0.877

*Note:*The significance of bold value is the best performers.

TABLE 4 Numbers of posts for each social support type after post classifications

Social support category	Initial posts	Percentage among all initial posts	Replies	Percentage among all replies
COM	4,719	5.05%	804,233	29.80%
SIS	78,399	83.89%	337,670	12.51%
SES	28,282	30.26%	234,810	8.70%
PIS	67,636	72.37%	962,542	35.66%
PES	6,259	6.70%	315,830	11.70%

length = 300 performs the best. Table 3 summarizes classification results. Among the six algorithms, XGBoost performs the best for COM, PES, and PIS classifications, while Logistic Regression is the best performer to classify SES and SIS. Hence, the best-performing algorithm for each social support category was then used to classify all the other posts that have not been annotated. Again, because each of the five social support classifier works independently, one post can belong to more than one social support category.

Similar to many other online communities, threaded discussions in the OHC we study have two types of posts: initial posts and replies. An "initial post" (or original post) starts a threaded discussion, which can be followed by zero or more "replies" (or comments). In our dataset, the total number of initial posts is 93,453. Among them, 83.9% are classified as SIS and 30.3% are classified as SES. The total number of replies is 2,699,144, with 35.7% classified as PIS, and 11.7% as PES. Classification results of unannotated posts are summarized in Table 4.

3.2 | Variables

Because this study investigates if users' first-time support-seeking experience can predict their subsequent participation in OHCs, we focused on a subset of users (referred to as "focal users") whose first posts initiated a thread to seek informational or emotional support (or both) in the OHC. In other words, for our analysis, each focal user corresponds to one focal thread (i.e., the thread started by the focal user's first support-seeking post) in the OHC. Then we used a regression model to reveal how their experience in such support-seeking threads are related to their subsequent engagement in the OHC.

Independent variables for our model measure the support a focal user received from her initial support-seeking post (i.e., thread). We measured the received support from two perspectives: (a) the quantity of received support and (b) the quality of received support. All independent variables are based on support received within 1 week after focal users' first support-seeking post.

1. *The quantity of received support* is defined as the number of replies in the thread started by the initial post. Note that we excluded self-replies, which were published by the user who started the thread. A larger number of replies to a thread indicates that the focal user who started the thread received more social support from the OHC.

2. *The quality of received support* is measured from two aspects. At the lexical level, we calculated *the length of replies*—the average number of characters in replies to the initial post. Lengthier replies are assumed to provide better support. At the semantic level, we measured the *match of support*—the extent to which a reply provides the type of support that was sought by the initial post. For example, if an initial post is SIS, a matching reply should have the PIS label. As a post can have multiple labels, an initial post that is both SIS and SES would be matched by a reply that is either PIS or PES (or both). For each thread in our pool, we calculated the percentage of replies providing support that matches the type of support sought by the initial post (e.g., PIS matches SIS and PES matches SES). A higher percentage indicates that replies from other users provide support that better matches what the support seeker sought in the initial post. Again, self-replies were excluded from the calculation.

The dependent variable in our model measures whether a user is still actively posting in the OHC one week after her first support-seeking initial post (posted at day *t*). While a user can still be involved in an OHC by lurking, posting represents a higher level of engagement that benefits other users and the whole community. It is also possible, although unlikely in an established OHC, that a small group of users keep posting content that is detrimental to the community. Also, in a moderated OHC, like the one we study, such users have been banned by the community with their content removed from our dataset.

To ensure the robustness of our results, we defined the dependent variables based on four different observation windows one week after *t*. Specifically, we measured if a focal user posted anything besides within her first thread during four different time periods from $t + 7$ to $t + k$, where k=29, 59, 89, and 179 days, respectively. If a focal user was active in posting during $t + 7$ to $t + k$ in the OHC, she was labeled as a positive instance (i.e., Class 1). Otherwise, she was labeled as Class 0.

Control variables include how active the user was during her first week *(#posts)*, the length of the initial post (*#initialPostLen*, measured by the number of characters), and if the thread still attracts replies after the first week *(ongoing)*. The first two control variables measure a user's intrinsic motivation to participate in the OHC. We also included *ongoing* because if a thread still attracted replies one week after its inception, it is more likely that the focal user who started the thread would come back to the OHC and post.

Therefore, our logistic regression model can be represented as follows:

$$\log\left(\frac{P(\mathrm{DV}=1)}{1-P(\mathrm{DV}=1)}\right)=\beta_0+\beta_1\cdot\mathrm{posts}$$
$$+\beta_2\cdot\mathrm{initialPostLen}+\beta_3\cdot\mathrm{ongoing}+\beta_4\cdot\mathrm{replies}$$
$$+\beta_5\cdot\mathrm{lengthReplies}+\beta_6\cdot\%\mathrm{SupportMatch}+\epsilon$$

where β_1, β_2, ..., β_6 are the coefficients, β_0 is a constant, and ϵ is the error term.

4 | RESULTS

Our user pool consisted of 17,547 focal users whose first post in the OHC was an initial post that sought informational or emotional support (or both). Among them, 34.5% kept posting from the second week to the first month (*k = 29*), 39.3 and 41.1% continued posting from the second week until the second month (*k = 59*) and the third month (*k = 89*), respectively. Last, 43.2% of focal users kept posting throughout the first 6 months (*k = 179*). Also, 485 users' initial posts received no replies during their first week. For them, all three independent variables *#replies*, *lengthReplies*, and *%SupportMatch* are set to 0 s. Table 5 lists all variables in our model and their summary statistics, and Table 6 reports their correlations. Because the distributions of *#replies*, *lengthReplies*, *#posts*, and *#initialPostLen* are highly skewed, we used log-transformed values of these variables in our analysis. Then, we standardized all control variables and independent variables with *Z*-scores, which helps us better observe how the changes on independent variables affect the users' continuous participation.

To address possible multicollinearity issues, we ran logistic regression using maximum likelihood estimation and checked variance inflation factors (VIFs). All VIF values were no larger than 2, suggesting that multicollinearity was not a concern. Also, as shown in Table 7, no variables are highly correlated with each other. We assessed model fit using Akaike information criterion (AIC) reported in Table 7.

Table 7 shows results of models with *DV-1*, *DV-2*, *DV-3*, and *DV-4* as dependent variables, respectively. All four models yield consistent findings: OHC users' first-time support-seeking experience matter for their continued participation in the OHC. *First*, adding variables of the support received by seekers can help to better predict seekers' subsequent engagement in the OHC as the AIC decreases from 19,190 in Model 0 to 18,878 in Model 1. *Second*, the quantity of received support, *#replies*, is a consistent, positive and significant predictor for future OHC engagement. *Third*, the quality of received support is also important for future engagement—a higher level

TABLE 5 Variable descriptions and summary statistics ($N = 17{,}547$)

Variable	Description	Mean	*SD*	(min, max)
DV-1	Continued posting during [$t + 7$, $t + 29$]	0.35	0.48	(0, 1)
DV-2	Continued posting during [$t + 7$, $t + 59$]	0.39	0.49	(0, 1)
DV-3	Continued posting during [$t + 7$, $t + 89$]	0.41	0.49	(0, 1)
DV-4	Continued posting during [$t + 7$, $t + 179$]	0.43	0.50	(0, 1)
#posts	The number of posts by the focal user during her first week, including other initial posts and replies to other threads	1.01	2.90	(0, 210)
#initialPostLen	The number of characters in the focal initial post by the focal user	954.49	778.93	(1, 13,745)
Ongoing	Whether there is still any non-self reply to the thread 1 week after its inception	0.33	0.47	(0, 1)
#replies	The total number of non-self-replies to the focal thread during the first week	5.26	6.14	(0, 258)
lengthReplies	The average number of characters of replies during a thread's first week	590.31	378.62	(0, 6,684.5)
%SupportMatch	The percentage of social support match between the initiated thread and non-self-replies during the first week	0.66	0.34	(0, 1)

TABLE 6 Pearson correlation coefficients among variables ($N = 17{,}547$)

		Correlations					
Variables		5	6	7	8	9	10
1	DV-1	0.35	0.10	0.26	0.21	0.07	0.05
2	DV-2	0.35	0.11	0.26	0.21	0.07	0.05
3	DV-3	0.34	0.10	0.26	0.21	0.07	0.05
4	DV-4	0.33	0.10	0.26	0.22	0.07	0.05
5	Log (#posts)	1.00	0.04	0.05	0.11	0.02	0.02
6	Log (#initialPostLen)		1.00	0.03	0.13	0.16	0.12
7	Ongoing			1.00	0.17	−0.12	−0.04
8	Log (#replies)				1.00	0.40	0.11
9	Log (lengthReplies)					1.00	0.60
10	%SupportMatch						1.00

of match between social support sought in an initial post and supported provided by replies promotes a seeker's subsequent posting activity. The length of replies received is also a positive predictor of support seekers' engagement, although it is only marginally significant.

As one would expect, all three control variables, *#posts*, *#initialPostLen*, and *ongoing*, have significant and positive effects on users' subsequent participation. A focal user who published more posts or published longer posts during her first week is more likely to post again later. In addition, if a focal user's first thread was still active one week after its inception, the focal user is more likely to be active in posting after the first week, perhaps because the focal user is still actively reading new replies and interacting with those who provided support in the thread.

In addition, AICs in Table 7 monotonically increase from Model 1 to Model 4. This suggests that it becomes more difficult to predict user's posting behavior further in the future, if such predictions are only based on users' activities and social support received during their first week. That said, whereas the strength of *#replies* as a predictor of continued participation remained basically the same across time, the predictive power of *%SupportMatch* increased over time. This pattern of results speaks to the value of matching support or minimizing support gaps as predictors of future participation in an OHC.

TABLE 7 Results of logistic regression models with different DVs

	Model 0	Model 1	Model 2	Model 3	Model 4
DV	DV-1	DV-1	DV-2	DV-3	DV-4
Intercept					
	−0.716*** (0.0179)	−0.727*** (0.0181)	−0.492*** (0.0175)	−0.397*** (0.0173)	−0.291*** (0.0171)
Control variables					
Log (#posts)	0.799*** (0.0188)	0.787*** (0.0190)	0.783*** (0.0192)	0.782*** (0.0193)	0.769*** (0.0194)
Log (#initialPostLen)	0.230*** (0.0187)	0.181*** (0.0191)	0.178*** (0.0186)	0.174*** (0.0184)	0.157*** (0.0181)
Ongoing	0.585*** (0.0172)	0.554*** (0.0179)	0.543*** (0.0176)	0.535*** (0.0175)	0.525*** (0.0175)
Independent variables					
Log (#replies)	-	0.282*** (0.0203)	0.292*** (0.0200)	0.312*** (0.0199)	0.323*** (0.0199)
Log (lengthReplies)	-	0.047. (0.0262)	0.048. (0.0253)	0.024 (0.0249)	0.011 (0.0246)
%SupportMatch	-	0.059* (0.0239)	0.061** (0.0230)	0.066** (0.0227)	0.085*** (0.0224)
AIC	19,190	18,878	19,668	19,925	20,236

Note: "***" $p < .001$, "**" $p < .01$, "*" $p < .05$, "." $p < .1$. *SE*s are in parentheses.

5 | CONCLUSION

This study examined how OHC users' early experience in seeking and receiving social support predict their subsequent community engagement. By detecting the seeking and providing different types of social support from user-generated content with text mining, this work represents the first to evaluate both the quantity and quality of received social support. In particular, enabled by large-scale machine-learning-based classification of support seeking and provision within posts, we proposed a novel measure for the quality of received social support—the match between social support sought and received. We also used regression models to investigate if users' first-time support-seeking experience can predict their future posting behavior in an OHC. Our analyses find the evidence that users' first-time support seeking experience, especially the quantity and the match of support they received from the community have positive effects on their subsequent participation.

Outcomes of our study can provide insight into the management and design of sustainable OHCs. For example, it is important to pay attention to new users who join an OHC to seek support. Their support-seeking posts can be prioritized and recommended to other users who have rich experience in providing the type of social support being sought. The importance of social support match is also highlighted by our finding that among replies to an initial post, an average of only 66% provided the type support sought by the original poster (see Table 5). By increasing the chance that a first-time support seeker receives enough of the type of support that is sought, an OHC can better retain users, which is a key factor for the success of OHCs. A supportive and sustainable OHC can benefit everyone involved in the community as well as future users who share similar health concerns. Meanwhile, we would encourage OHCs to consider their users' privacy concerns and ethical implications when they deploy any intervention based on our findings.

This research also has limitations. First, we only studied users who started their posting behavior with an initial post to seek social support. However, some new users may seek support, especially informational support, by lurking instead of posting. Without users' clickstream data, it is difficult to investigate such lurking support seekers' behavior. Second, our explanatory models only reveal the correlation between users' first-time experience seeking support with subsequent participation. Thus, although the causality between our independent and dependent variables would make sense, a rigorous inference of causal relationships needs more work to handle endogeneity issues (e.g., a user's offline health status). Third, our research investigated users' first-time support seeking experience, while their later experience in exchanging support and interacting with others can also be important predictors for their future engagement. For example, an interesting direction for future work is to increase the length of observation window so that one can use longer time series or sliding time windows to predict users' future engagement as in (Wang et al., 2020). A longer observation period would also enable researchers to use the Granger causality approach to reveal how users' earlier engagement casually affects their later engagement (Gopalsamy, Semenov, Pasiliao, McIntosh, & Nikolaev, 2017). In addition, we did not differentiate between different types of support match, while some match may be more important to user's participation than others. Another limitation of our study is that none of the research team members belong to this OHC. Therefore, our analyses of content (e.g., annotating post and developing machine learning algorithms) shared by these OHC members may lead to misrepresentations because we may lack the specific knowledge and experience of OHC members (Andalibi, Haimson, De Choudhury, & Forte, 2018). One possible way to address this issue is to include community members into our team for future investigations. Last but not the least, we only worked with data from one OHC for breast cancer and future work is needed to generalize our findings to other OHCs.

We also acknowledge that there are ethical implications to conducting research involving automated data collection from the web or other means of unobtrusive sampling, despite the public-facing nature of the OHC analyzed in this study and its statement notifying users of their potential inclusion in research. There are admittedly gray areas in terms of what is or is not appropriate when users do not consent to participating in a particular study. Although the support group used in this study clearly indicates that users might be included in research, many users do not read privacy policies and some people have difficulty understanding them (Obar & Oeldorf-Hirsch, 2020). Issues of consent might also evolve given changes in technology and societal thought about privacy. The OHC analyzed in this study has existed for decades, and users posting in the early 2000s may not have anticipated the large-scale data scraping methods utilized in this study. Based on these concerns, some scholars highlight issues of privacy and confidentiality along with informed consent as the foremost ethical considerations for conducting research using techniques of big data in the context of health (Ienca et al., 2018). Moreover, it is unclear when this OHC added its privacy policy indicating the potential for research. Although the policy is meant to be inclusive, if it was added after 2013 (the last date included in the data for this project), users would have posted prior to reading that statement. More generally, users' expectations regarding the messages they post on an OHC need to be more thoroughly considered.

Scholars underscore the importance of protecting subjects when analyzing large-scale online data (Vitak, Shilton, & Ashktorab, 2016). Therefore, we took users' privacy seriously. By automatically coding posts and manually rewording posts (e.g., in Table 1), we did not use users' own words in our manuscript, so that it becomes more difficult for readers to search a quote online to identify users. Still, if privacy policies and terms of service, including those that mention the potential for research are regarded as "the biggest lie on the internet" (Obar & Oeldorf-Hirsch, 2018), researchers need to thoroughly consider the ethical implications of data scraping and the necessity of being transparent and honest with participants, just as they would in a traditional study. Specifically, for this study, the chance of any harm to the participants is quite small. We hope that the potential benefits of learning how to sustain user participation in OHCs outweighs the small chance of harm.

ORCID

Xi Wang https://orcid.org/0000-0002-7568-0851
Kang Zhao https://orcid.org/0000-0002-8321-2804

ENDNOTES

[1] https://www.breastcancer.org/about_us/bco_commitment/privacy_statement

[2] We listed one example post for each category of social support. To preserve users' privacy, we replaced some texts with dots and reworded some sentences from these posts.

REFERENCES

Andalibi, N., Haimson, O. L., De Choudhury, M., & Forte, A. (2018). Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *25*(5), 1–35.

Arora, N. K., Finney Rutten, L. J., Gustafson, D. H., Moser, R., & Hawkins, R. P. (2007). Perceived helpfulness and impact of social support provided by family, friends, and health care providers to women newly diagnosed with breast cancer. *Psycho-Oncology*, *16*(5), 474–486.

Bambina, A. (2007). *Online social support: The interplay of social networks and computer-mediated communication*, Youngstown, NY: Cambria Press.

Barak, A., Boniel-Nissim, M., & Suler, J. (2008). Fostering empowerment in online support groups. *Computers in Human Behavior*, *24*(5), 1867–1883.

Barbee, A. P., & Cunningham, M. R. (1995). An experimental approach to social support communications: Interactive coping in close relationships. *Annals of the International Communication Association*, *18*(1), 381–413.

Caplan, R. D. (1987). Person-environment fit theory and organizations: Commensurate dimensions, time perspectives, and mechanisms. *Journal of Vocational Behavior*, *31*(3), 248–267.

Caplan, S. E., & Turner, J. S. (2007). Bringing theory to research on computer-mediated comforting communication. *Computers in Human Behavior*, *23*(2), 985–998.

Chancellor, S., Hu, A., & De Choudhury, M. (2018). Norms matter: Contrasting social support around behavior change in online weight loss communities. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. pp. 1–14. https://doi.org/10.1145/3173574.3174240

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chou, W.-Y. S., Hunt, Y. M., Beckjord, E. B., Moser, R. P., & Hesse, B. W. (2009). Social media use in the United States: Implications for health communication. *Journal of Medical Internet Research*, *11*(4), e48.

Chuang, K. Y., & Yang, C. C. (2014). Informational support exchanges using different computer-mediated communication formats in a social media alcoholism community. *Journal of the Association for Information Science and Technology*, *65*(1), 37–52.

Cobb, N. K., Graham, A. L., & Abrams, D. B. (2010). Social network structure of a large online community for smoking cessation. *American Journal of Public Health*, *100*(7), 1282–1289.

Cobb, S. (1976). Social support as a moderator of life stress. *Psychosomatic Medicine*, *38*, 300–314.

Crowley, J. L., & High, A. C. (2020). Effects of provider identity and locus of supportive conversation on coping with an identity-threatening stressor. *Communication Research*, *47*(4), 487–511.

Cutrona, C. E., Shaffer, P. A., Wesner, K. A., & Gardner, K. A. (2007). Optimally matching support and perceived spousal sensitivity. *Journal of Family Psychology*, *21*(4), 754–758.

Cutrona, C. E., & Suhr, J. A. (1992). Controllability of stressful events and satisfaction with spouse support behaviors. *Communication Research*, *19*(2), 154–174.

De Choudhury, M., & De, S. (2014). *Mental health discourse on reddit: Self-disclosure, social support, and anonymity*. Paper presented at Eighth International AAAI Conference on Weblogs and Social Media.

Eaker, E. D. (2005). Social support and physical health: Understanding the health consequences of relationships. *American Journal of Epidemiology*. *161*(3), 297–298. https://doi.org/10.1093/aje/kwi036.

Fox, S. (2014). *The social life of health information*. Retrieved from https://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/

Ganz, P. A., Desmond, K. A., Leedham, B., Rowland, J. H., Meyerowitz, B. E., & Belin, T. R. (2002). Quality of life in long-term, disease-free survivors of breast cancer: A follow-up study. *Journal of the National Cancer Institute*, *94*(1), 39–49.

Goldsmith, D. J. (1994). The role of facework in supportive communication. In *Communication of social support: Messages, interactions, relationships, and community* (pp. 29–49). Thousand Oaks, CA: SAGE Publications.

Gopalsamy, R., Semenov, A., Pasiliao, E., McIntosh, S., & Nikolaev, A. (2017). Engagement as a driver of growth of online health forums: Observational study. *Journal of Medical Internet Research*, *19*(8), e304.

Graham, A. L., Zhao, K., Papandonatos, G. D., Erar, B., Wang, X., Amato, M. S., … Pearson, J. L. (2017). A prospective examination of online social network dynamics and smoking cessation. *PLoS One*, *12*(8), e0183655.

Hawn, C. (2009). Take two aspirin and tweet me in the morning: How twitter, Facebook, and other social media are reshaping health care. *Health Affairs*, *28*(2), 361–368.

Healey, B., Hoek, J., & Edwards, R. (2014). Posting behaviour patterns in an online smoking cessation social network: Implications for intervention design and development. *PLoS One*, *9*(9), e106603.

Helgeson, V. S., Cohen, S., Schulz, R., & Yasko, J. (2000). Group support interventions for women with breast cancer: Who benefits from what? *Health Psychology*, *19*(2), 107–114.

High, A. C., Jennings-Kelsall, V., Solomon, D. H., & Marshall, A. D. (2015). Military families online: Seeking and providing support through internet discussion boards. In *A communication perspective on the military: Interactions, messages, and discourses* (pp. 101–120). New York, NY: Peter Lang.

High, A. C., & Crowley, J. L. (2018). Gaps among desired, sought, and received support: Deficits and surpluses in support when coping with taboo marital stressors. *Communication Research*, *45*(3), 319–338.

High, A. C., & Steuber, K. R. (2014). An examination of support (in) adequacy: Types, sources, and consequences of social support among infertile women. *Communication Monographs*, *81* (2), 157–178.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining*. pp. 168–177.

Hwang, K. O., Ottenbacher, A. J., Green, A. P., Cannon-Diehl, M. R., Richardson, O., Bernstam, E. V., & Thomas, E. J. (2010). Social support in an internet weight loss community. *International Journal of Medical Informatics*, *79*(1), 5–13.

Ienca, M., Ferretti, A., Hurst, S., Puhan, M., Lovis, C., & Vayena, E. (2018). Considerations for ethics review of big data health research: A scoping review. *PLoS One*, *13*(10), e0204937.

Jacobson, D. E. (1986). Types and timing of social support. *Journal of Health and Social Behavior*, *27*, 250–264.

Jones, R., Sharkey, S., Smithson, J., Ford, T., Emmens, T., Hewis, E., … Owens, C. (2011). Using metrics to describe the participative stances of members within discussion forums. *Journal of Medical Internet Research*, *13*(1), e3.

Katz, J. E., Rice, R. E., & Aspden, P. (2001). The internet, 1995-2000: Access, civic involvement, and social interaction. *American Behavioral Scientist*, *45*(3), 405–419.

Khanpour, H., Caragea, C., & Biyani, P. (2018). *Identifying emotional support in online health communities*. Paper presented at Thirty-Second AAAI Conference on Artificial Intelligence.

Kim, E., Han, J. Y., Moon, T. J., Shaw, B., Shah, D. V., McTavish, F. M., & Gustafson, D. H. (2012). The process and effect of supportive message expression and reception in online breast cancer support groups. *Psycho-Oncology*, *21*(5), 531–540.

Linden, W., & Vodermaier, A. (2012). Mismatch of desired versus perceived social support and associated levels of anxiety and depression in newly diagnosed cancer patients. *Supportive Care in Cancer*, *20*(7), 1449–1456.

Lv, Y., Ng, W. W. Y., Lee, J. W. T., Sun, B., & Yeung, D. S. (2008). *Information extraction based on information fusion from multiple news sources from the web*. Paper presented at 2008 IEEE International Conference on Systems, Man and Cybernetics. pp. 1471–1476.

Ma, H., Smith, C. E., He, L., Narayanan, S., Giaquinto, R. A., Evans, R., Hanson, L., & Yarosh, S. (2017). Write for life: Persisting in online health communities through expressive writing and social support. *Proceedings of the ACM on human-computer interaction, 1(CSCW)*. pp. 1–24. https://doi.org/10.1145/3134708

MacGeorge, E. L., Feng, B., & Burleson, B. R. (2011). Supportive communication. In *Handbook of interpersonal communication* (Vol. 4, pp. 317–354). Thousand Oaks, CA: Sage.

Matire, L. M., Stephens, M. A. P., Druley, J. A., & Wojno, W. C. (2002). Negative reactions to received spousal care: Predictors and consequences of miscarried support. *Health Psychology*, *21* (2), 167–176.

Maunsell, E., Brisson, J., & Deschěnes, L. (1995). Social support and survival among women with breast cancer. *Cancer*, *76*(4), 631–637.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *ArXiv Preprint ArXiv*, *1301*, 3781 Retrieved from http://arxiv.org/abs/1301.3781

Namkoong, K., Shah, D. V., Han, J. Y., Kim, S. C., Yoo, W., Fan, D., … Gustafson, D. H. (2010). Expression and reception of treatment information in breast cancer support groups: How health self-efficacy moderates effects on emotional well-being. *Patient Education and Counseling*, *81*, S41–S47.

Obar, J. A., & Oeldorf-Hirsch, A. (2018). The clickwrap: A political economic mechanism for manufacturing consent on social media. *Social Media + Society*, *4*(3), 2056305118784770.

Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, *23*(1), 128–147.

Qiu, B., Zhao, K., Mitra, P., Wu, D., Caragea, C., Yen, J., Greer, G. E., & Portier, K. (2011). *Get online support, feel better—Sentiment analysis and dynamics in an online cancer survivor community*. Paper presented at 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing. pp. 274–281.

Rains, S. A., Peterson, E. B., & Wright, K. B. (2015). Communicating social support in computer-mediated contexts: A meta-analytic review of content analyses examining support messages shared online among individuals coping with illness. *Communication Monographs*, *82*(4), 403–430.

Reynolds, J. S., & Perrin, N. A. (2004). Mismatches in social support and psychosocial adjustment to breast cancer. *Health Psychology*, *23*(4), 425–430.

Shumaker, S. A., & Brownell, A. (1984). Toward a theory of social support: Closing conceptual gaps. *Journal of Social Issues*, *40* (4), 11–36.

Siewert, K., Antoniw, K., Kubiak, T., & Weber, H. (2011). The more the better? The relationship between mismatches in social support and subjective well-being in daily life. *Journal of Health Psychology*, *16*(4), 621–631.

Sudau, F., Friede, T., Grabowski, J., Koschack, J., Makedonski, P., & Himmel, W. (2014). Sources of information and behavioral patterns in online health forums: Observational study. *Journal of Medical Internet Research*, *16*(1), e10.

Thong, M. S. Y., Kaptein, A. A., Krediet, R. T., Boeschoten, E. W., & Dekker, F. W. (2007). Social support predicts survival in dialysis patients. *Nephrology Dialysis Transplantation*, *22*(3), 845–850.

Vitak, J., Shilton, K., & Ashktorab, Z. (2016). Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. pp. 941–953.

Vlahovic, T. A., Wang, Y.-C., Kraut, R. E., & Levine, J. M. (2014). Support matching and satisfaction in an online breast cancer support community. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 1625–1634.

Walther, J. B., & Boyd, S. (2002). Attraction to computer-mediated social support. *Communication Technology and Society: Audience Adoption and Uses*, *153188*, 50–88.

Wang, Xi., Zhao, K., & Street, N. (2014). *Social support and user engagement in online health communities*. In International Conference on Smart Health. pp. 97–110. Cham: Springer. https://doi.org/10.1007/978-3-319-08416-9_10

Wang, X., Zhao, K., & Street, N. (2017). Analyzing and predicting user participations in online health communities: A social support perspective. *Journal of Medical Internet Research*, *19*(4), e130. https://doi.org/10.2196/jmir.6834

Wang, Xi, Zuo, Z., & Zhao, K. (2015). *The evolution and diffusion of user roles in online health communities: A social support perspective*. 2015 International Conference on Healthcare Informatics. pp. 48–56.

Wang, X., Zhao, K., Zhou, X., & Street, N. (2020). Predicting user posting activities in online health communities with deep learning. *ACM Transactions on Management Information Systems (TMIS)*, *11*(3), 1–15. https://doi.org/10.1145/3383780

Wang, Y.-C., Kraut, R., & Levine, J. M. (2012). To stay or leave? The relationship of emotional and informational support to commitment in online health support groups. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. pp. 833–842.

Wicks, P., Massagli, M., Frost, J., Brownstein, C., Okun, S., Vaughan, T., … Heywood, J. (2010). Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Internet Research*, *12*(2), e19.

Xu, R., Zhou, J., Zhang, Q., & Hendler, J. A. (2018). *Research on Online Health Communities, A Systematic Review. Encyclopedia of Social Network Analysis and Mining* (2nd edition). New York, NY: Springer. https://doi.org/10.1007/978-1-4939-7131-2

Yang, D., Kraut, R., & Levine, J. M. (2017). Commitment of newcomers and old-timers to online health support communities. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. pp. 6363–6375.

Yoo, W., Namkoong, K., Choi, M., Shah, D. V., Tsang, S., Hong, Y., … Gustafson, D. H. (2014). Giving and receiving emotional support online: Communication competence as a moderator of psychosocial benefits for women with breast cancer. *Computers in Human Behavior*, *30*, 13–22.

Zhang, M., & Yang, C. C. (2017). Post classification and recommendation for an online smoking cessation community. In *Public Health Intelligence and the Internet* (pp. 59–73). Cham: Springer. https://doi.org/10.1007/978-3-319-68604-2_4.

Zhang, S., Grave, E., Sklar, E., & Elhadad, N. (2017). Longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks. *Journal of Biomedical Informatics*, *69*, 1–9.

Zhao, K., Wang, X., Cha, S., Cohn, A. M., Papandonatos, G. D., Amato, M. S., … Graham, A. L. (2016). A multirelational social network analysis of an online health community for smoking cessation. *Journal of Medical Internet Research*, *18*(8), e233. https://doi.org/10.2196/jmir.5985

How to cite this article: Wang X, High A, Wang X, Zhao K. Predicting users' continued engagement in online health communities from the quantity and quality of received support. *J Assoc Inf Sci Technol*. 2021;72:710–722. https://doi.org/10.1002/asi.24436