RESEARCH

Assessing the needs of patients with breast cancer and their families across various treatment phases using a Latent Dirichlet Allocation model: a text-mining approach to online health communities

Chaojin Da[1] · Yiwen Duan[2] · Zhenying Ji[1] · Jialin Chen[2] · Haozhi Xia[2] · Yajuan Weng[2] · Tingting Zhou[2] · Changrong Yuan[2] · Tingting Cai[2]

Abstract

Purpose This study aimed to assess the different needs of patients with breast cancer and their families in online health communities at different treatment phases using a Latent Dirichlet Allocation (LDA) model.

Methods Using Python, breast cancer-related posts were collected from two online health communities: patient-to-patient and patient-to-doctor. After data cleaning, eligible posts were categorized based on the treatment phase. Subsequently, an LDA model identifying the distinct need-related topics for each phase of treatment, including data preprocessing and LDA topic modeling, was established. Additionally, the demographic and interactive features of the posts were manually analyzed.

Results We collected 84,043 posts, of which 9504 posts were included after data cleaning. Early diagnosis and rehabilitation treatment phases had the highest and lowest number of posts, respectively.

LDA identified 11 topics: three in the initial diagnosis phase and two in each of the remaining treatment phases. The topics included disease outcomes, diagnosis analysis, treatment information, and emotional support in the initial diagnosis phase; surgical options and outcomes, postoperative care, and treatment planning in the perioperative treatment phase; treatment options and costs, side effects management, and disease prognosis assessment in the non-operative treatment phase; diagnosis and treatment options, disease prognosis, and emotional support in the relapse and metastasis treatment phase; and follow-up and recurrence concerns, physical symptoms, and lifestyle adjustments in the rehabilitation treatment phase.

Conclusion The needs of patients with breast cancer and their families differ across various phases of cancer therapy. Therefore, specific information or emotional assistance should be tailored to each phase of treatment based on the unique needs of patients and their families.

Keywords Latent Dirichlet Allocation · Breast cancer · Online health communities · Cancer therapy

Abbreviations
LDA   Latent Dirichlet Allocation
OHCs Online health communities

Changrong Yuan and Tingting Cai contributed to the work equally.

✉ Changrong Yuan
yuancr@fudan.edu.cn

✉ Tingting Cai
caitingtingguo@163.com

Chaojin Da
Dachaojin16@163.com

Yiwen Duan
23211170002@m.fudan.edu.cn

Zhenying Ji
jizhenying.666@163.com

Jialin Chen
chenjl604@163.com

Haozhi Xia
mickey_sn15@163.com

Yajuan Weng
faith830406@hotmail.com

Tingting Zhou
21211170008@m.fudan.edu.cn

1 Department of Nursing, School of Clinical Nursing, Gansu Health Vocational College, Lanzhou, China

2 School of Nursing, Fudan University, 305 Fenglin Road, Shanghai 200032, China

Background

According to the GLOBOCAN 2020 estimates, female breast cancer is the most common type of cancer, with an estimated 2.3 million new cases, and is the fifth leading cause of cancer-related deaths, accounting for 685,000 deaths [1]. The disease burden of breast cancer varies significantly among countries, with China topping the list of the number of cases and death rate [2]. With the development of medical technology and popularization of breast cancer screening programs, early diagnosis and treatment of breast cancer have progressed significantly, reducing mortality [3, 4]. Patients with breast cancer often experience multiple concurrent and interacting symptoms throughout their cancer journey, which require obtaining relevant health information to cope with the disease [5]. Patients and their families often have diverse needs at different treatment stages [6]. Among these, information and emotional needs are the most common. In a cross-sectional study, patients with breast cancer in China scored low on instrumental, emotional, and informational support [7]. Several family members of patients with breast cancer reportedly scored moderately high on information needs and low on emotional function and mental health [8, 9]. However, the existing research is mainly cross-sectional surveys and semi-structured interviews, which are often limited to samples from specific geographic areas, thus restricting the widespread applicability of the findings. In addition, an in-depth exploration of the needs of patients and their families during the various treatment stages is lacking.

With the rapid development of internet technology and widespread use of online social platforms, an increasing number of patients and their families seek online support [10]. Patients with breast cancer and their families reportedly often search for cancer-related issues on internet platforms to increase their understanding of the disease and achieve personal education goals. Additionally, they quickly obtain and share disease-related information through online message platforms and support groups, providing emotional support to each other, reducing anxiety and loneliness, and promoting positive disease coping mechanisms [11]. Therefore, the analysis and mining of data from social platforms are necessary to understand the health information needs of patients with breast cancer.

Approximately 420,000 new cases of breast cancer are diagnosed in China annually. Recently, the incidence of breast cancer has increased by 3–4% annually. The peak age of onset of female breast cancer is 45–55 years in China, which is earlier than 55–59 years in Western countries. The early onset implies a heavier disease burden on the patients, their families, and society [12]. In China, breast cancer is a high-frequency search topic in online health communities (OHCs). Patients with breast cancer and their families often discuss disease-related issues, share medical experiences, exchange information, and express fear and frustration with others by posting messages on OHCs [13]. These posts are genuine logs of self-reported needs, feelings, and knowledge shared by patients with breast cancer and their families from different regions of the country during various treatment phases that are regularly updated [13]. The evolving and multifaceted needs of patients with breast cancer and their families throughout their treatment journey can be understood by text-mining these narratives. In traditional Chinese culture, breast cancer is considered an emotional issue or the result of karmic retribution, leading to feelings of stigma in patients and even vicarious stigma of their caregivers, causing them to be reluctant to seek help from their families and professionals [7]. Although cross-sectional and longitudinal analyses of topics discussed on Chinese breast cancer forums have been performed, in-depth exploration of the needs of patients with breast cancer and their families is lacking, and few studies have explored the needs of patients with breast cancer by text-mining Chinese social media posts.

OHC-related data are enormous, with complex structures and diverse content, making manual analysis and processing inefficient and ineffective. To extract valuable information, using natural language processing and machine learning technologies for automated text analysis is necessary. To determine the most suitable topic modeling method for this study, we conducted a thorough evaluation of several mainstream topic modeling techniques at the initial stage of the research, covering the advanced BERT-based (Bidirectional Encoder Representations from Transformers) BERTopic model, traditional Latent Semantic Indexing (LSI), and Latent Dirichlet Allocation (LDA) [14–16]. Through comparative analysis, we noted that the unsupervised learning model, LDA, is particularly effective in mining deep themes in texts manifested through lexical co-occurrence. Compared to BERTopic and LSI, LDA shows significant advantages in topic consistency and its performance stability under different parameter configurations, highlighting its uniqueness in the field of topic modeling. Moreover, an increasing number of researchers are using the LDA model to analyze the health information needs of internet users, suggesting that this method has unique advantages in text mining for web-based question-and-answer communities [17, 18]. Hence, this study plans to use the LDA model to assess the needs of breast cancer patients and their families in online health communities, in order to provide a scientific basis for formulating targeted nursing measures. Due to the argumentation process's complexity, a comprehensive explanation is provided in the Electronic Supplementary Material of the paper.

Methods

A five-step process, including data source determination, collection, cleaning and coding, preprocessing, and LDA topic modeling was included in the LDA model. The overall process is illustrated in Fig. 1.

Data source determination

Initially, health-related websites were searched using the Baidu search engine using "breast cancer" as the main keyword. After a detailed examination of the websites, breast cancer-related posts appeared mainly in two OHCs: patient-to-patient and patient-to-doctor. Patient-to-patient communities are an open platform where patients with breast cancer and their families provide peer support and exchange tips and information on how to cope with cancer. "Breast Bar," a sub-social forum of Baidu Post Bar, is the largest (approximately 237,181 members) and most active (approximately 1,573,096 posts) patient-to-patient community for breast cancer in China. Patient-to-doctor communities are internet-based platforms that enable patients to seek medical consultations from doctors. "Xunyiwenyao," "Haodaifu," and "360Liangyi" are all significant and highly influential large-scale physician–patient interaction

Fig. 1 Flowchart of data processing and modeling
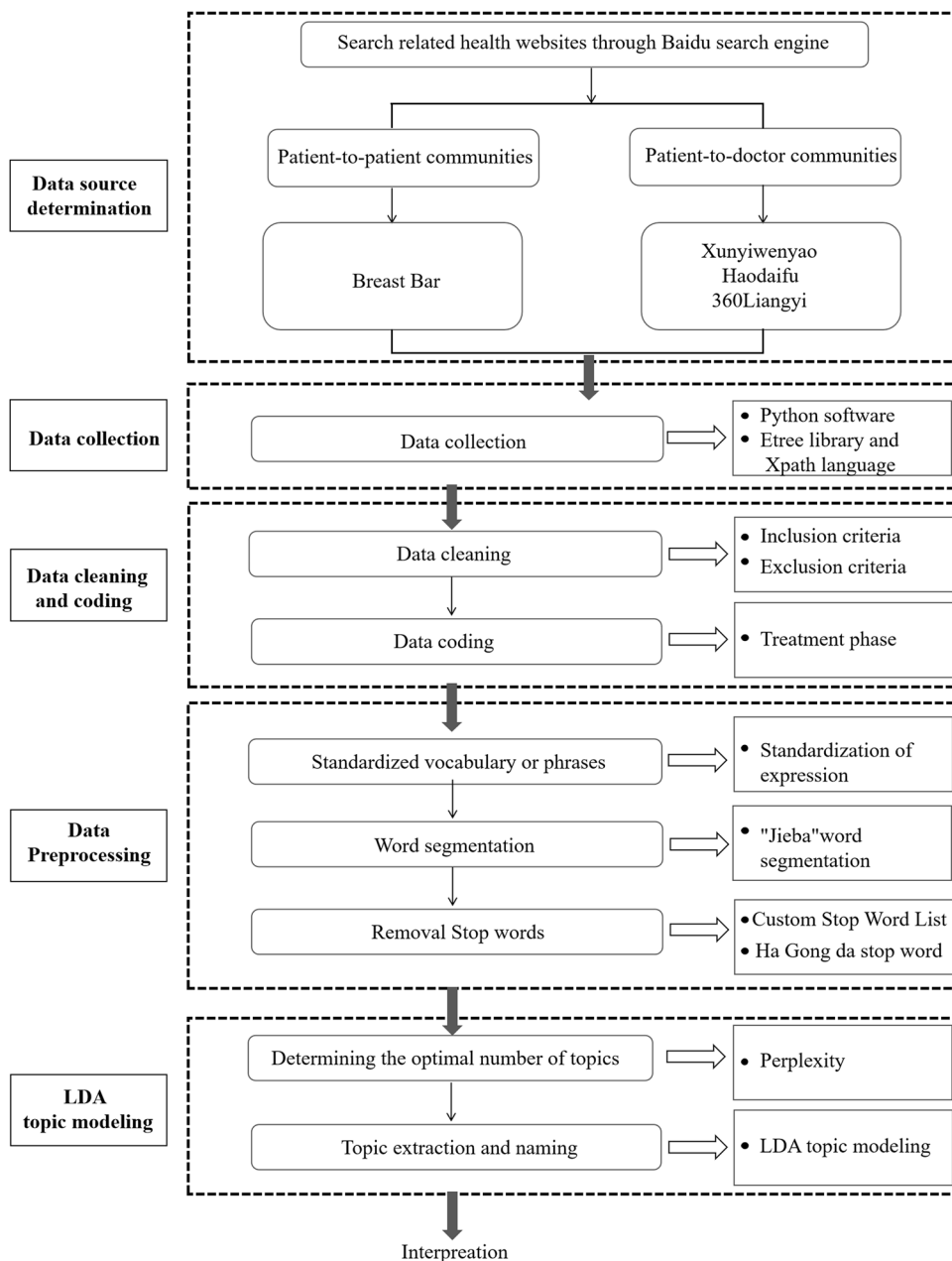
platforms in China. These platforms have > 900,000 doctors from > 100,000 legitimate hospitals nationwide. Therefore, we selected "Breast Bar," "Xunyiwenyao," "Haodaifu," and "360Liangyi" as data sources.

Data collection

Using the keyword "breast cancer," we applied the Python software (Python Software Foundation, CWI, Netherlands) to collect posts published in the "Breast Bar," "Xunyiwenyao," "Haodaifu," and "360Liangyi" from January 1, 2017, to April 25, 2023. Data scraping and storage were achieved by utilizing Python and its library resources, such as Requests and lxml. The core process of scraping primarily relied on the Requests library for making GET requests and parsing HTML content to acquire data. For web pages with complex structures, advanced parsing tools were employed or user actions simulated to precisely extract the needed information. Additionally, the speed of scraping was enhanced through the use of multithreading and asynchronous programming techniques. Throughout the development of the web scraper, we adhered to websites' privacy policies, only collecting data necessary for the research, paid attention to personal data protection, and avoided unnecessary server load by setting appropriate request headers, avoiding peak hours, and implementing suitable delays. The Etree library and Xpath language were used to filter advertising posts during the data scraping process on the aforementioned forum and websites. Available post-related information, including user ID, post title and body content, timestamp, number of likes, number of favorites, and number of comments, was stored in a database.

Data cleaning and coding

To improve the veracity of the dataset and resolve any disagreements, the first author and a research assistant independently evaluated, filtered, and coded the text data. Inclusion criteria included (1) posts on breast cancer-related needs, (2) posts from which the current treatment stage of a patient can be inferred, and (3) posts published by the patients or their families. Exclusion criteria included (1) posts with repeated or duplicated questions and (2) posts that were irrelevant to the research question. The posts were classified based on the treatment phase, including initial diagnosis, perioperative, non-operative, relapse and metastasis, and rehabilitation treatment phases.

Data preprocessing

We preprocessed the extracted posts using the Python software. First, as the informality and colloquial nature of the language used in online communication often result in the use of different sentences with similar meanings, which might impede natural language processing, we initially standardized vocabulary or phrases that convey identical meanings but vary in wording. For instance, "单子" (examination form) and "报告单" (report sheet) were unified as "报告单" (report sheet). Next, texts were segmented into words before further analysis. We used the extensively used Chinese word segmentation tool, "Jieba" (结巴) library for this purpose. After text segmentation, both the Ha Gong da and customized stop word lists were used to eliminate punctuation marks, conjunctions, numbers, English letters, and a set of frequently used words with minimal meaningful content. Following this preprocessing step, meaningful noun phrases were extracted.

LDA model

During the LDA topic modeling phase, the Gensim library was utilized, which is a Python library specifically designed for topic modeling, document indexing, and similarity retrieval of large corpora [19]. Using the LDA model, topics were extracted from collected posts, revealing latent topics hidden within the texts and differentiating the various meanings of words. The optimal number of topics was determined by adopting the perplexity calculation method, where a lower perplexity value indicates a better generalization capability of the model [17, 18]. After extracting the topics, each was characterized by a set of words with varying strengths, indicating the degree of association with the topic. Words with higher strength values are more representative of their respective topics. By analyzing the semantics of these topic words, we summarized the core content of each topic and assigned descriptive names to enhance understanding of their meanings.

When posting inquiries, posters often reveal their own or the patient's identity information as well as the patient's age. In this study, "poster identity" denotes clear information identifying the poster as either the patient or a relative. Posts lacking direct identification as such were excluded, except for age details, which were deemed non-essential for identity; absent age data led to classification under "unknown" without post exclusion. For example, a post might describe: "My mother, 56 years old, was recently diagnosed with triple-negative breast cancer. How long can someone with triple-negative breast cancer live?" Such descriptions enabled us to determine whether the poster is the patient themselves or a relative of the patient, and to ascertain the patient's age. Posts that do not disclose the patient's identity information are excluded based on our inclusion criteria. Ultimately, we compiled data on the number of posts made by the patients themselves, relatives of patients, patients under the age of 40, patients aged 40 and above, and posts from different

regions for each period. The region of the patients was inferred from the IP address captured during data collection.

Results

Basic information about the posts

We collected 84,043 posts, of which 9504 were included after data cleaning, including 7786 (81.92%) and 1718 (18.8%) from patient-to-patient and patient-to-doctor communities, respectively. Specific information collected from each community is presented in Table 1. After coding, 4902 posts represented the initial diagnosis phase, ranking first among the five treatment phases and accounting for 51.58% of the total posts. Conversely, the posts for the rehabilitation

phase were 588 (6.19%), which were the lowest among all treatment phases. The posts for perioperative, non-operative, and relapse and metastasis treatment phases were 1134, 1992, and 888 posts, respectively. The posts for the non-operative treatment phase had the highest average number of likes per post, with 10.92 likes per post. The average number of favorites marked and comments per post were 8.60 and 68.22 for the relapse and metastasis treatment phases, respectively, both ranking first. Detailed information on the average engagement metrics across the treatment phases is shown in Fig. 2.

Post-related demographic information

Only 23.60% of the posts contained information about the patient's age. During the perioperative phase, posts from young patients with breast cancer were higher than those from other older patients. However, in the other three treatment phases, the posts from the young patients were fewer than those from the older patients. In each treatment phase, the number of posts published by the patients exceeded those published by their family members. User location information was available for only 38.06% of posts. Throughout every treatment phase, East and Northwest China consistently generated the highest and lowest number of posts, respectively. Specific post-related demographic information is presented in Table 2.

Table 1 Basic information on post-collection

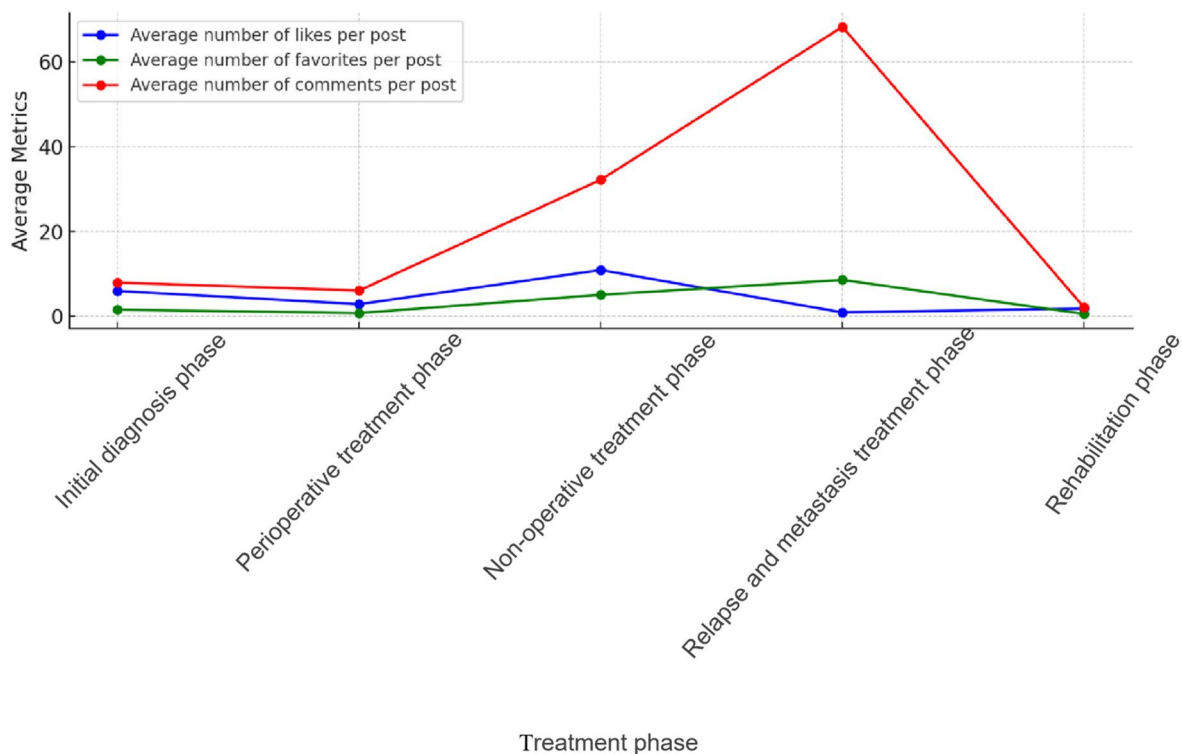| Online health communities	Name of online health communities	Number of raw posts	Number of valid posts
Patient-to-patient communities	Breast Bar	81,693	7786
Patient-to-doctor communities	Xunyiwenyao	748	615
	Haodaifu	1200	898
	360Liangyi	316	205

Fig. 2 Average engagement metrics across treatment phases

Table 2 Post-related demographic information (*n*, %)

Item	Category	Initial diagnosis phase (*n*=4902)	Perioperative treatment phase (*n*=1134)	Non-operative treatment phase (*n*=1992)	Relapse and metastasis treatment phase (*n*=888)	Rehabilitation phase (*n*=588)	Total (*n*=9504)
Patient age	<40	532 (10.85)	168 (14.81%)	61 (3.06%)	90 (10.14%)	122 (20.75%)	973 (10.24%)
	≥40	567 (11.57%)	150 (13.23%)	101 (5.07%)	252 (28.38%)	200 (34.01%)	1270 (13.36%)
	Unknown	3803 (77.5%)	816 (71.96%)	1830 (91.87%)	546 (61.49%)	266 (45.24%)	7261 (76.40)
User identity	Patient himself	4012(81.84%)	931(80.51%)	1322(66.35%)	537(60.53%)	355(60.41%)	727(76.56%)
	Patient's family	890(18.16%)	221(19.49%)	670(33.65%)	351(39.47%)	233(39.59%)	222(23.44%)
User location	Northeast China	70 (1.43%)	81 (7.14%)	104 (5.22%)	42 (4.73%)	6 (1.02%)	303 (3.19%)
	North China	292 (5.96%)	69 (6.08%)	286 (14.36%)	102 (11.49%)	66 (11.22%)	815 (8.58%)
	Northwest China	63 (1.29%)	19 (1.68%)	6 (0.30%)	42 (4.73%)	0 (0.00%)	130 (1.37%)
	Central China	51 (1.04%)	50 (4.41%)	67 (3.36%)	18 (2.03%)	54 (9.18%)	240 (2.53%)
	Southwest China	171 (3.49%)	31 (2.73%)	24 (1.20%)	54 (6.08%)	30 (5.10%)	310 (3.26%)
	East China	545 (11.12%)	131 (11.55%)	439 (22.04%)	180 (20.27%)	108 (18.37%)	1403 (14.76%)
	South China	190 (3.88%)	50 (4.41%)	104 (5.22%)	48 (5.41%)	24 (4.08%)	416 (4.38%)
	Unknown	3520 (71.81%)	703 (61.99%)	962 (48.29%)	402 (45.27%)	300 (51.02%)	5887 (61.94%)

Analysis of need-related topics among patients with breast cancer and their family members

To determine the optimal number of topics, we calculated the perplexity of each topic. Figure 3 shows a line graph of the perplexity for each treatment phase, in which the horizontal and vertical axes represent the number of optional topics and perplexity value, respectively. We chose a topic number in the interval $N \in [1, 10]$ to ensure that valuable topics would be discovered and each topic would be well explained. The perplexity values for each treatment phase first decreased rapidly as the number of topics increased, reaching their lowest point at three topics for the early diagnosis stages and two topics each for the perioperative, non-operative, relapse and metastasis, and rehabilitation phases. Subsequently, as the number of topics increased, the perplexity values gradually increased. Therefore, with three topics for the initial diagnosis phase and two for the other phases, the model was the most stable, and the intelligibility of the topic words was the best. Table 3 shows the 11 final topics with their respective topic words and strengths. These topics enabled us to explain and summarize the discussion text for each topic.

Initial diagnosis phase

Three topics were identified in the initial diagnosis phase, including disease outcomes, diagnostic analysis, treatment information, and emotional support.

Disease outcomes (Topic 1)

Keywords such as "cure rate, life expectancy, prognosis, recurrence, and survival period" suggest that the primary concerns of individuals newly diagnosed with breast cancer and their families are the prospects of curing the disease and improving their overall survival. For example, similar to some other posts, one patient asked, "I am 32 years old, and today I received a pathological diagnosis of invasive breast cancer. How long can I live?".

Diagnosis analysis (Topic 2)

Keywords related to auxiliary examinations were included, such as "immunohistochemistry, puncture, ultrasound (B-mode), biopsy, color Doppler ultrasound, magnetic resonance imaging(MRI), and report sheet," indicating the attitude of patients and their families in seeking assistance from doctors or fellow patients on online platforms to interpret medical reports, thus obtaining comprehensive diagnostic information. A patient asked, "I underwent a color ultrasound and needle biopsy at the Beijing Cancer University Hospital 8 days ago, and the diagnosis was breast cancer. Can anyone please help me review my report? Is this type relatively less malignant?".

Treatment information and emotional support (Topic 3)

Keywords such as "treatment method, chemoradiotherapy, surgery, breast-conservation, targeted therapy, costs, and hospitals" showed that patients with breast cancer and their families sought information about treatment plans, costs, and hospitals. For instance, in one post, an individual asked for help because the doctor mentioned three treatment plans for a mother's invasive breast cancer diagnosis. Keywords such as "breakdown, fear, and positive energy" indicated that patients and families often

Fig. 3 Perplexity change graph. Initial diagnosis phase (**A**), perioperative treatment phase (**B**), non-operative treatment phase (**C**), relapse and metastasis treatment phase (**D**), and rehabilitation phase (**E**)
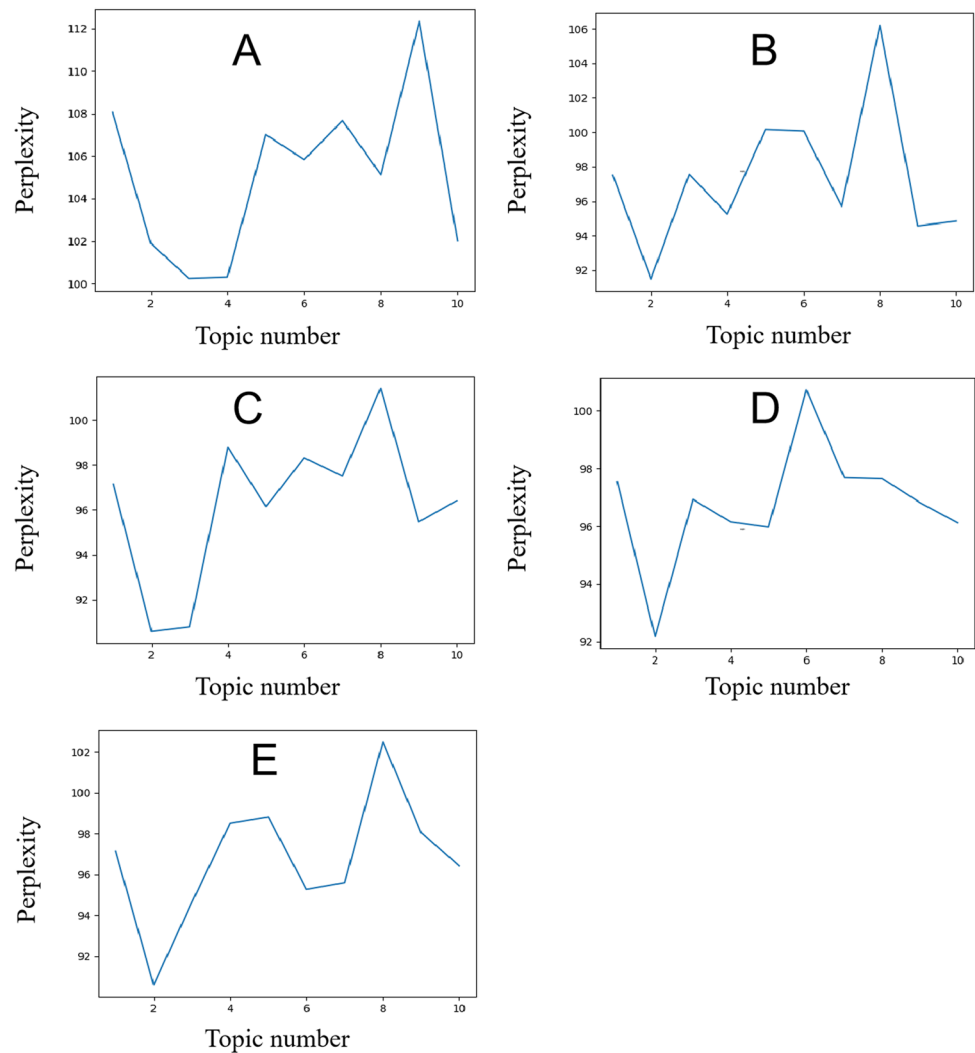
experience complex emotions, including "emotional turmoil, desperation, sadness, and helplessness," following a breast cancer diagnosis. They expressed their feelings online and sought support and confidence by sharing their experiences. For example, a newly diagnosed patient with breast cancer expressed feelings of fear and helplessness, asking for support and sharing experiences from others, to which other users responded.

Perioperative treatment phase

The perioperative treatment phase included two topics: surgical options and outcomes and post-surgical care and treatment planning.

Surgical options and outcomes (Topic 4)

Keywords such as "recurrence, reconstruction, prognosis, artificial breast, implantation, modified radical surgery,

breast-conserving surgery, total resection, infection, and general anesthesia" indicated that patients and their families were concerned about breast cancer surgery and its outcomes. Surgery is the primary treatment for early- and mid-stage breast cancer. Surgical options include modified radical mastectomy, breast-conserving surgery, and breast reconstruction after mastectomy. Patients could choose the most suitable option based on their situation. However, navigating these choices can be daunting; therefore, patients often seek help from healthcare professionals or peers for more information to aid their decision-making. Common questions included "We need to choose a surgical plan. Have any patient had their whole breasts removed or received implants? If so, can you share what it was like for them?".

Postoperative care and treatment planning (Topic 5)

The keywords in this topic were related to post-surgical care and follow-up treatment. Keywords such as "diet, recovery

Table 3 Distribution and strength of topics and keywords among patients with breast cancer and families across treatment phases

Treatment phase	Topic no	Topic label	Topic strength	Keywords (strength)
Initial diagnosis phase	1	Disease outcomes	0.663	Cure rate (0.131); life expectancy (0.123); carcinoma in situ (0.098); prognosis (0.089); recurrence (0.064); probability of cure (0.043); life span (0.042); survival period (0.035); middle-late Stage (0.024); seriousness (0.014)
	2	Diagnosis analysis	0.585	Immunohistochemistry (0.138); puncture (0.104); triple-negative (0.064); ultrasound (B-mode) (0.048); biopsy (0.047); triple-positive (0.042); color Doppler ultrasound (0.042); magnetic resonance imaging (MRI) (0.038); stage (0.033); report sheet (0.029)
	3	Treatment information and emotional support	0.393	Therapeutic method (0.059); breakdown (0.047); chemoradiotherapy (0.043); surgery (0.040); breast-conservation (0.040); targeted therapy (0.038); fear (0.038); positive energy (0.033); cost (0.031); hospital (0.024)
Perioperative treatment phase	4	Surgical options and outcomes	0.962	Recurrence (0.179); reconstruction (0.148); prognosis (0.122); artificial breast (0.100); implantation (0.085); modified radical surgery (0.068); breast-conserving (0.053); total resection (0.047); infection (0.047); general anesthesia (0.036)
	5	Postoperative care and treatment planning	0.678	Diet (0.181); drainage (0.129); chemotherapy (0.085); pain (0.066); recovery time (0.063); functional exercise (0.063); swelling (0.041); radiotherapy (0.021); weight (0.015); endocrine therapy (0.014)
Non-operative treatment phase	6	Treatment options and costs	0.955	Targeted therapy (0.224); course of treatment (0.141); paclitaxel (0.130); doxorubicin (0.130); medical insurance (0.120); Adriamycin and Cyclophosphamide Regimen (0.056); Taxotere (docetaxel) and Cyclophosphamide Regimen (0.047); drug change (0.037); treatment plan (0.037); expense (0.033)
	7	Side effects management and disease prognosis assessment	0.930	Recrudescence (0.283); tamoxifen (0.141); endometrium (0.116); albumin (0.099); nausea (0.090); peripherally inserted central catheters (PICCs) (0.058); diet (0.053); prognosis (0.047); skin ulcers (0.037); hair loss (0.006)

Table 3 (continued)

Treatment phase	Topic no	Topic label	Topic strength	Keywords (strength)
Relapse and metastasis treatment phase	8	Diagnosis and treatment options	0.915	Nodules (0.265); computed tomography (0.229); hydrops (0.093); ultrasound (0.060); report sheet (0.048); endocrine therapy (0.044); bone scan (0.049); radiotherapy (0.044); liver (0.044); chemotherapy (0.039)
	9	Disease prognosis and emotional support	0.674	Survival time (0.232); stress (0.106); life span (0.072); targeted therapy (0.063); narrowing (0.043); sorrow (0.038); psychological preparation (0.034); communication (0.033); despair (0.027); encouragement (0.026)
Rehabilitation treatment phase	10	Follow-up and recurrence concerns	0.843	Lymph (0.228); blood flow (0.148); scanning (0.113); computed tomography (0.104); B-ultrasound (0.076); cancer antigen 153 (0.076); recurrence (0.073); color ultrasound (0.013); centimeter (0.007); armpit (0.005)
	11	Physical symptoms and lifestyle adjustments	0.694	Edema (0.148); arm (0.141); pain (0.141); diet (0.127); moxibustion (0.090); weight-bearing (0.013); traditional Chinese medicine (0.010); intumescence (0.010); exercise (0.007); sexual life (0.007)

period, functional exercise, and weight" indicated that patients and their families wanted information on diet, exercises, and activities after breast cancer surgery. Keywords such as "drainage, pain, and swelling" reflected their interest in drainage tube maintenance and common postoperative symptom management. For example, the question, "Has anyone in this group undergone breast reconstruction surgery? If so, I am curious to know how long the drainage tube needs to be kept in place." Additionally, "chemotherapy and radiotherapy" indicated concerns about future treatment plans after breast cancer. For instance, similar to some posts, one user asked, "My wife underwent breast resection for ductal carcinoma in situ approximately ten days ago. Does the patient require chemotherapy?".

Non-operative treatment phase

The non-operative treatment phase included two topics: treatment options and costs and management of chemoradiotherapy side effects and disease prognosis assessment.

Treatment options and costs (Topic 6)

Keywords including "targeted therapy, course of treatment, paclitaxel, doxorubicin, medical insurance, Adriamycin and Cyclophosphamide Regimen, Taxotere (docetaxel) and Cyclophosphamide Regimen, drug change, treatment plan, and expense" indicated that during the non-operative treatment phase, patients and their families often have questions about treatment plans, including the effectiveness and cost of chemotherapy drugs. Common questions included, "The doctor at the city hospital suggested the Adriamycin and Cyclophosphamide Regimen, while the doctor at the provincial hospital recommended the Taxotere (docetaxel) and Cyclophosphamide Regimen. Which one is better?" "The doctor recommended changing the medication to either docetaxel or albumin-bound paclitaxel. Which should we choose?" "Is doxorubicin covered by insurance, and how much would an additional eight chemotherapy treatments cost?".

Side effects management and disease prognosis assessment (Topic 7)

This topic focused on the side effects of chemoradiotherapy and disease prognosis. Discussions on this topic were centered on the effective management of common side effects, such as "endometrial thickening, albumin decline, nausea, skin ulcers, and hair loss," and disease prognosis after non-surgical treatment. Common questions included, "I have been on targeted therapy for almost 2 months, and my white blood cell count is low. What should I eat to help them?

Has anyone experienced taxophene-induced endometrial thickening, and how did you deal with it? I have completed chemoradiotherapy for breast cancer, and the doctor said that it went well. Is there a high chance of recurrence?" Additionally, as patients often need peripherally inserted central catheters (PICCs) for multiple chemotherapy sessions, "PICC" was reflected in this topic. Patients and their families sought information on the maintenance of PICCs. For example, one patient asked, "My arm, where the PICC was, has been swollen and painful for three days. What should I do?".

Relapse and metastasis treatment phase

The relapse and metastasis treatment phases included diagnosis and treatment options, disease prognosis, and emotional support.

Diagnosis and treatment options (Topic 8)

This topic focused on the diagnosis and treatment of breast cancer recurrence and metastasis with keywords including "nodules, computed tomography, hydrops, B-ultrasound, report sheet, bone scan, and liver." Patients and their families often share diagnostic reports on the platform to determine whether cancer has recurred or spread. A common question was, "How can I tell if there is bone metastasis?" When facing recurrence or metastasis, they sought advice on treatment options, such as asking, "Can radiotherapy control brain metastases from breast cancer in my mother?".

Disease prognosis and emotional support (Topic 9)

This topic included keywords such as "survival time, stress, life span, targeted therapy, narrowing, sorrow, psychological preparation, communication, despair, and encouragement," mainly addressing disease prognosis and emotional support. Patients and their families often share feelings of pessimism and despair after learning about cancer recurrence and metastasis. They sought to understand their expected survival time, emotional support, and encouragement from online platforms. For example, "My breast cancer has recurred! It is overwhelming, and I do not know how long I have, but I am not giving up. I need encouragement!" Many patients shared their feelings and received supportive responses.

Rehabilitation treatment phase

The rehabilitation phase included follow-up and recurrence concerns, physical symptoms, and lifestyle adjustments.

Follow-up and recurrence concerns (Topic 10)

Keywords such as "lymph, blood flow, scanning, computed tomography, B-ultrasound, cancer antigen 153, color ultrasound, centimeter, and armpit" represented the posts related to review reports for patients with breast cancer who need regular check-ups. They often share their results on online platforms, asking for feedback. For example, "My mother had breast cancer surgery 4 years ago; is this review report good?" The keyword "recurrence" showed that patients and their families are still concerned about cancer recurrence, even in the rehabilitation treatment phase, as seen in questions like "I have had breast cancer for 7 years; could it come back?".

Physical symptoms and lifestyle adjustments (Topic 11)

Keywords such as "pain and edema" showed that patients with breast cancer and their families desire nursing knowledge regarding common physical symptoms during recovery. Patients often struggle with limb pain and swelling after breast cancer treatment. For example, "I finished my breast cancer treatment 1 year ago, but my right arm still hurts and swells. What can I do?" Keywords such as "diet, moxibustion, weight-bearing, traditional Chinese medicine, exercise, and sexual life" involved questions on diet, exercise, sexual activity, and using traditional Chinese medicine in recovery. People often asked "Is it safe to have sex 2 years after breast cancer surgery on my right side? Can it cause a recurrence?" or "I am 30 and finished treatment 6 months ago. What should be eaten to prevent a recurrence? Are there Chinese medicinal treatments to help?".

Discussion

To the best of our knowledge, this is the first study to use the LDA model to analyze posts in two OHCs in China to evaluate the needs of patients with breast cancer and their families at different treatment stages. The findings of this study indicated that the number of posts gradually decreased as therapy progressed, with the highest and lowest number of posts observed in the early diagnosis and rehabilitation treatment phases, which was consistent with the results of Mikal et al. [20]. In the early diagnosis phases, patients and their families typically possessed limited knowledge about the disease and were eager to obtain information [21]. However, as treatment progressed, patients developed a deeper understanding of the disease, and simultaneously, the sociocultural and psychological barriers they faced gradually diminished, which can partly explain the declining trend in the number

of online posts throughout therapy [22]. Furthermore, the mismatch between the support available online and the needs of patients may have contributed to fewer posts. Mikal et al. studied online social support among patients with breast cancer and found that patients and their families primarily post on Facebook to seek resource support; however, they often received emotional support instead because it is cost-effective, reducing participation by patients and their families on social media [20].

Regarding post-interaction, Mikal et al. revealed that post-engagement metrics showed a sharp increase after cancer diagnosis, followed by a steady tapering that continued throughout the transition to cancer therapy [20]. Our data showed a similar trend, with likes peaking during the non-operative treatment phase, whereas favorites and comments peaked during the relapse and metastasis treatment phase. This indicates that posts during the non-operative and recurrence and metastasis treatment phases generated higher engagement and resonance among other users, resulting in more interaction. In the context of the unequal distribution of medical resources in China, patients and their families are likely to encounter common problems, such as the side effects of radiotherapy, chemotherapy, and neoadjuvant therapy during non-operative treatment [23]. They often lack offline support for these common issues and are more inclined to seek information online, where they can find resonance with others. During the relapse and metastasis treatment phase, patients and their families often experience significant psychological pressure. However, Chinese have roundabout and restrained emotional expression; therefore, they are reluctant to share their emotions with close individuals [24]. In contrast, they prefer to use the Internet to express their concerns and seek relevant information. The posts, which document the emotional journeys of patients and their families with cancer, such as fear, hope, resilience, and love, had a profound impact on readers, easily evoking deep emotional connections. Moreover, these posts offer valuable information and emotional support to individuals facing similar circumstances because stories of personal struggle and resilience could be inspiring for other patients, building a virtual support system [25].

Notably, our findings revealed that the majority of posts were from patients with breast cancer, which aligned with the relatively young age profile of Chinese patients with breast cancer. Breast cancer in several Asian countries primarily occurs at the ages of 45–69 years, which is > 10 years earlier than that in Europe and the USA [26]. These patients generally possessed proficient internet skills and the ability to independently seek breast cancer-related information online. Notably, East China contributed the highest number of posts, potentially because of its superior economic development and greater accessibility to Internet resources [27].

Regarding the post contents, patients with breast cancer and their family members predominantly sought information and emotional support through online media rather than other sources. This finding was consistent with the results of Nadine et al., who showed that emotional and information support were among the most frequently provided or received types of support in interpersonal interactions as well as computer-mediated environments for individuals with cancer or other serious illnesses [28]. The results highlighted a greater emphasis on treatment-related information within the post content, such as attending hospitals and chemotherapy regimens, compared with nursing-related information, which was consistent with the findings of the study by Bei on the information needs of Hong Kong Chinese patients with breast cancer [29]. This could be attributed to social media users being mostly patients or family members who possess general knowledge about treatment-related aspects but might lack expertise in highly specialized disease nursing information [30]. Moreover, this finding corresponds with the results of Cho, who reported that visitation rates for treatment-related issues in breast cancer surpassed those for other categories [31].

Our study found that patients with breast cancer and their family members consistently expressed ongoing concerns about cancer recurrence, metastasis, and overall survival duration throughout the entire disease trajectory, which is consistent with the results of the study by Shi [32]. This emotional state, characterized by "fear, worry, or concern about cancer returning or progressing," is commonly known as fear of cancer recurrence (FCR) [33]. Herschbach et al. confirmed that FCR might appear at any stage of the illness and persist for many years even after treatment or remission of the illness, which is consistent with our research findings [34]. This is primarily because cancer progression is uncertain, and patients and family members must constantly seek information to manage changes in symptoms, drug resistance, and complications and to understand how to cope with the negative emotions that accompany uncertainty in cancer progression, helping them have a sense of control over their lives [35–37].

Miller found that cancer-related demands varied across a patient's cancer trajectory [38]. Similarly, we discovered that patients and their families had different needs during the different stages of treatment. During the initial diagnostic stage, patients with breast cancer and their family members were more concerned about disease outcomes, diagnostic analysis, and treatment information. They experienced strong anxiety regarding the types of hospital examinations and diagnostic results, as well as treatment methods and their effectiveness. During the perioperative period, patients and their families focused on surgical choices, prognosis, postoperative care, and follow-up treatments, particularly information related to surgical procedures, as keywords such as "reconstruction, implantation,

modified radical mastectomy, breast conservation, and total mastectomy," which were related to surgical methods, repeatedly appeared with a high degree of strength. Breast cancer surgery, especially mastectomy, has a profound impact on the appearance of patients and significantly affects their emotional well-being and social functioning [39, 40]. Therefore, patients and their families frequently aim to alleviate the impact on their bodies through breast conservation or reconstruction while concurrently harboring concerns about the potential adverse consequences of these surgeries, prompting active seeking of relevant information online [41]. During the non-operative treatment period, treatment options and costs, chemoradiotherapy side effect management, and disease prognosis assessment were the main concerns of patients and their family members. This might be because patients at this stage had a heavy economic burden after a period of treatment, and some side effects of radiotherapy and chemotherapy emerged and progressed as the treatment progressed. Patients with recurrent or metastatic breast cancer and their family members should pay more attention to the auxiliary diagnosis, treatment, and prognosis. Recurrence and metastasis are the main causes of death in patients with breast cancer. In patients with breast cancer, this is a recurring negative event [42]. Consequently, patients might seek a second medical opinion to ensure the accuracy of the diagnosis and re-evaluate treatment options in the relapse and metastasis treatment phase. During the rehabilitation treatment phase, patients and their families showed heightened interest in understanding follow-up reports, receiving detailed explanations of prognosis, acquiring knowledge about managing physical symptoms effectively, and maintaining healthy lifestyles. After treatment is completed, patients enter a long-term monitoring and follow-up phase. In contrast, breast cancer survivors usually experience treatment-related lymphedema, pain, and fatigue [43, 44]. We found that Chinese breast cancer survivors turned to traditional Chinese medicines, such as moxibustion and herbal medicine, to alleviate specific post-treatment discomfort and enhance their overall quality of life. An exploratory study of the main French forums and discussion groups indicated that herbal medicines are frequently cited for breast cancer [45]. Breast cancer survivors expressed a strong desire for knowledge related to their sex lives. Sex was not a priority during treatment compared to securing survival; however, as time progressed, the significance of sex and intimacy in their lives became more evident, playing a substantial role in the overall quality of life [46]. Notably, some patients with breast cancer have misconceptions, believing that engaging in sexual activity might increase the risk of breast cancer recurrence. Therefore, healthcare providers should offer comprehensive sex education to breast cancer survivors. In conclusion, the patients and their families had a strong initiative to seek information at each stage. One possible reason why patients and their families resort to online queries is the increasing complexity of treatments involving multiple healthcare providers in different locations, which leads to a lack of treatment continuity

[47]. Additionally, some patients and their families might not have received sufficient information and emotional support from healthcare professionals, causing them to question treatment and care measures implemented by medical personnel, which motivated them to actively seek evidence to establish trust with healthcare providers involved in their treatment [48].

Our research findings revealed that although the emotional support needs of patients and their families were particularly pronounced during the early diagnosis and recurrent and metastatic treatment phases, the keywords associated with "emotional support needs" were not entirely consistent between these two phases, suggesting variations in the emotional support required by patients and their families during these different stages. During the early diagnosis phase, the keywords related to emotions were "breakdown" and "fear." A cancer diagnosis is often considered an extremely stressful and potentially traumatic experience [49]. Sussman et al. studied specialized oncological care following surgery for patients newly diagnosed with breast and colorectal cancer and revealed that the transition from cancer diagnosis to the formal cancer phase was accompanied by a surge in emotional distress, such as mental stress and anxiety, which correspondingly heightened the demand for information support and emotional care, which is consistent with our findings [50]. Therefore, when facing a breast cancer diagnosis, the primary feelings of the patients and their families are extreme helplessness, anxiety, and worry. In the recurrent and metastatic treatment phase, the keywords related to emotions were "pain," "psychological preparation," and "breakdown." Recurrence and metastasis are the primary causes of breast cancer-related deaths, representing a negative event for both patients and their families [42]. Consequently, patients and their families may be more prone to feelings of sadness and despair, even after being emotionally prepared for their impending death. A qualitative interview study conducted by Jassim et al. on 12 patients with breast cancer indicated that women diagnosed with late-stage disease were more likely to experience negative emotions [51]. However, both the early diagnosis and recurrent and metastatic treatment phases exhibited keywords related to positive emotions, such as "positivity" and "encouragement." This observation implies that despite the daunting challenges they faced, a number of patients and their families managed to sustain a sense of optimism, proactively confronting the recurrence or progression of the disease with resilience. Nevertheless, upon examining the frequency and prominence of the keywords, it became evident that patients and their families predominantly conveyed negative emotions on social platforms. Freedman et al. analyzed 1,024,041 posts related to breast cancer treatment and found that 57% of the posts expressed negative emotions, which is consistent with the findings of our study [22]. Positive psychological coping contributes to the emotional well-being of patients [22, 52]. Consequently, healthcare providers should greatly emphasize providing psychological support to patients during both the initial diagnosis and recurrent and metastatic treatment phases.

Study limitations

This study has some limitations. Although this study focused on the information needs of patients with breast cancer and their families at different stages of treatment, we did not specifically distinguish between the needs of the patients and their families. Furthermore, we did not conduct an in-depth comparison of the content of the posts from the two online health communities, where the needs of patients and their families may vary depending on whether they seek advice from doctors or fellow patients. Additionally, the information on the websites included in our study may have primarily reflected the perspectives of younger patients, potentially overlooking the comprehensive needs of older patients.

Conclusions

Our findings indicate that the information and emotional support needs of patients with breast cancer and their families differ across various phases of cancer therapy, and the need for emotional support intensifies in the early diagnostic and recurrence or metastasis phases. Concurrently, concerns regarding cancer recurrence, spread, and survival duration remain consistent throughout the therapeutic journey. Consequently, offering specific information or emotional assistance tailored to each phase of treatment is crucial for meeting the unique needs of patients and their families.

Author contribution Data analysis and interpretation, drafting of the manuscript, and final approval were performed by Tingting Cai, Changrong Yuan, and Chaojin Da. Study conception and design were performed by Yiwen Duan, Zhenying Ji, and Jialin Chen. The manuscript revision was performed by Haozhi Xia, Yajuan Weng, and Tingting Zhou. All authors participated fully in this work and assumed public responsibility for the relevant part of the content.

Declarations

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in this study.

Consent for publication All authors read the final version of the paper and approved its submission for publication.

Competing interests The authors declare no competing interests.

References

1. Sung H, Ferlay J, Siegel RL et al (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in185 countries. CA Cancer J Clin 71(3):209–249

2. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, Zeng H, Zhou J, Wei W (2021) Global patterns of breast cancer incidence and mortality: a population-based cancer registry data analysis from 2000 to 2020. Cancer Commun 41(11):1183–1194

3. Cai T, Zhou T, Chen J, Huang Q, Yuan C, Wu F (2023) Identification of age differences in cancer-related symptoms in women undergoing chemotherapy for breast cancer in China. BMC Women's Health 23(1):100

4. Cai T, Zhou T, Huang Q, Wu F, Ni F, Yuan C (2023) Cancer-related symptoms among young and middle-aged women undergoing chemotherapy for breast cancer: application of latent class analysis and network analysis. Eur J Oncol Nurs 63:102287

5. Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL (2015) Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. J Med Internet Res 17(4):e106

6. Occhiogrosso RH, Ren S, Tayob N, Li T (2022) Feasibility of an online patient community to support older women with newly diagnosed breast cancer. 22(2):178-185

7. Cai T, Huang Q, Yuan C (2021) Profiles of instrumental, emotional, and informational support in Chinese breast cancer patients undergoing chemotherapy: a latent class analysis. BMC Womens Health 21(1):183

8. Vahidi M, Mahdavi N, Asghari E, Ebrahimi H, Eivazi Ziaei J, Hosseinzadeh M, Namdar H, Kermani IA (2016) Other side of breast cancer: factors associated with caregiver burden. Asian Nurs Res 10(3):201–206

9. Harmon DM, Young CD, Bear MA, Aase LA, Pruthi S (2021) Integrating online community support into outpatient breast cancer care: Mayo Clinic Connect online platform. Digit Health 7:2055–2076

10. Basch CH, Sarah A, Phillip B, Charles E (2019) Readability of online breast cancer information. Breast J 25(3):562–563

11. Chen L, Wang P, Ma X, Wang X (2021) Cancer communication and user engagement on Chinese social media: content analysis and topic modeling study. J Med Internet Res 23(11):e26310

12. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K et al (2021) Global patterns of breast cancer incidence and mortality: a population-based cancer registry data analysis from 2000 to 2020. Cancer Commun 41(11):1183–1194

13. Zhang S, Grave E, Sklar E, Elhadad N (2017) Longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks. J Biomed Inform 69:1–9

14. Lee YJ, Park A, Roberge M, Donovan H (2022) What can social media tell us about patient symptoms: a text-mining approach to online ovarian cancer forum. Cancer Nurs 45:E27–E35

15. Grootendorst M (2022) BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv. Available online at: http://arxiv.org/abs/2203.05794 (accessed July, 2023)

16. Zhu G, Hao M, Zheng C, Wang L (2022) Design of knowledge graph retrieval system for legal and regulatory framework of multilevel latent semantic indexing. Comput Intell Neurosci 2022:6781043

17. Cabling ML, Turner JW, Hurtado-de-Mendoza A, Zhang Y, Jiang X, Drago F, Sheppard VB (2017) Sentiment analysis of an online breast cancer support group: communicating about tamoxifen. Health Commun 33(9):1158–1165

18. Hagg LJ, Merkouris SS, O'Dea GA, Francis LM, Greenwood CJ, Fuller-Tyszkiewicz M, Westrupp EM, Macdonald JA, Youssef GJ

(2022) Examining analytic practices in Latent Dirichlet Allocation within psychological science: scoping review. J Med Internet Res 24(11):e33166

19. Řehůřek R, Sojka P (2010) Software framework for topic modelling with large corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, pp. 45–50, Valletta, Malta: ELRA

20. Mikal JP, Beckstrand MJ, Grande SW et al (2021) Online support seeking and breast cancer patients: changes in support seeking behavior following diagnosis and transition off cancer therapy. Health Commun 36(6):731–740

21. Finney RLJ, Agunwamba AA, Wilson P et al (2016) Cancer-related information seeking among cancer survivors: trends over a decade (2003–2013). J Cancer Educ 31:348–357

22. Freedman RA, Viswanath K, Vaz-Luis I et al (2016) Learning from social media: utilizing advanced data extraction techniques to understand barriers to breast cancer treatment. Breast Cancer Res Tr 158:395–405

23. Lu X, Zhang R (2021) Impact of patient information behaviours in online health communities on patient compliance and the mediating role of patients' perceived empathy. Patient Educ and Couns 104(1):186–193

24. Sun WN, Hsu HT, Ko NY et al (2020) Decision-making processes in surrogates of cancer patients in a Taiwan intensive care unit. Int J Env Res Pub He 17(12):4443

25. Naganathan G, Bilgen I, Cleland J et al (2022) #COVID19 and #Breastcancer: a qualitative analysis of tweets. Curr Oncol 29(11):8483–8500

26. Ren W, Chen M, Qiao Y et al (2022) Global guidelines for breast cancer screening: a systematic review. Breast 64:85–99

27. Xu C, Wang Y, Yang H et al (2019) Association between cancer incidence and mortality in web-based data in China: infodemiology study. J Med Internet Res 21(1):e10677

28. Bol N, Rising CJ, Weert JCM (2022) Perceived emotional and informational support for cancer: patients' perspectives on interpersonal versus media sources. Communications 47(2):171–194

29. Bei AW, Lai MT, Choi KC et al (2015) Factors in the prioritization of information needs among Hong Kong Chinese breast cancer patients. Asia-Pac J Oncol Nurs 2(3):176–185

30. Ohigashi A, Ahmed S, Afzal AR et al (2017) Breast cancer information communicated on a public online platform: an analysis of 'Yahoo! Answer Japan' J Prim Health Care 9(2):167–172

31. Cho H, Silver N, Na K et al (2018) Visual cancer communication on social media: an examination of content and effects of #Melanomasucks. J Med Internet Res 20(9):e10501

32. Shi Y, Dai J, Kuai B et al (2023) Impact of fear of recurrence in patients with cancer on caregivers' psychological health: a meta-analysis. Asia-Pac J Oncol Nurs 10(10):100299

33. Lebel S, Ozakinci G, Humphris G et al (2016) From normal response to clinical problem: definition and clinical features of fear of cancer recurrence. Support Care Cancer 24:3265–3268

34. Herschbach P, Berg P, Dankert A et al (2005) Fear of progression in chronic diseases: psychometric properties of the fear of progression questionnaire. J Psychosom Res 58(6):505–511

35. Kállay É, Medrea F, Dégi CL (2022) On top of that all, now Covid-19, too. A scoping review of specificities and correlates of fear of cancer recurrence in breast cancer patients during COVID-19. Breast 62:123–134

36. Schook RM, Linssen C, Schramel FMNH et al (2014) Why do patients and caregivers seek answers from the Internet and online lung specialists? A qualitative study. J Med Int Res 16(2):e2842

37. Browne S, Dowie AL, Mitchell ED et al (2011) Patients' needs following colorectal cancer diagnosis: where does primary care fit in? Brit J Gen Pract 61(592):e692–e699

38. Miller LE (2014) Uncertainty management and information seeking in cancer survivorship. Health Commun 29(3):233–243

39. Skórzewska M, Kurylcio A, Rawicz-Pruszyński K et al (2021) Impact of mastectomy on body image and sexuality from a LGBTQ perspective: a narrative review. J Clin Med 10(4):567

40. Fortunato L, Loreti A, Cortese G et al (2021) Regret and quality of life after mastectomy with or without reconstruction. Clin Breast Cancer 21(3):162–169

41. Shekhar N, Jaiswal R, Joseph L et al (2023) An overview of psychological analysis of breast cancer patients undergoing modified radical mastectomy and breast conservation surgery and its impact on objectified body consciousness at a Tertiary Care Cancer Centre in South India. Clin Breast Cancer 23(7):e394–e400

42. Wang X, Wang N, Zhong L et al (2020) Prognostic value of depression and anxiety on breast cancer recurrence and mortality: a systematic review and meta-analysis of 282,203 patients. Mol Psychiatr 25(12):3186–3197

43. Sullivan CW, Leutwyler H, Dunn LB et al (2018) Stability of symptom clusters in patients with breast cancer receiving chemotherapy. J Pain Symptom Manag 55(1):39–55

44. Yoo TK, Jang M, Lee E et al (2018) Endocrine treatment-related symptoms and patient outcomes in breast cancer: a meta-analysis. J Breast Cancer 21(1):37–44

45. Lognos B, Carbonnel F, Launay IB et al (2019) Complementary and alternative medicine in patients with breast cancer: exploratory study of social network forum data. JMIR Cancer 5(2):e12536

46. Ghizzani A, Bruni S, Luisi S (2018) The sex life of women surviving breast cancer. Gynecol Endocrinol 34(10):821–825

47. Horner K, Ludman EJ, McCorkle R et al (2013) An oncology nurse navigator program designed to eliminate gaps in early cancer care. Clin J Oncol Nurs 17(1):40–48

48. Bickell NA, LePar F, Wang JJ et al (2007) Lost opportunities: physicians' reasons and disparities in breast cancer treatment. J Clin Oncol 25(18):2516–2521

49. Costanzo ES, Ryff CD, Singer BH (2009) Psychosocial adjustment among cancer survivors: findings from a national survey of health and well-being. Health Psychol 28(2):147

50. Sussman J, Bainbridge D, Whelan TJ et al (2018) Evaluation of a specialized oncology nursing supportive care intervention in newly diagnosed breast and colorectal cancer patients following surgery: a cluster randomized trial. Support Care Cancer 26:1533–1541

51. Jassim GA, Whitford DL (2014) Understanding the experiences and quality of life issues of Bahraini women with breast cancer. Soc Sci Med 107:189–195

52. Gall TL, Bilodeau C (2017) "Why me?"–women's use of spiritual causal attributions in making sense of breast cancer. Psych Heal 32(6):709–727