Analysis on childhood fever health information seeking behaviors in online health community using a mixed-methods approach

Xuan Pan[1,2] (iD), Zhuoyuan Tang[3], Ying Liu[1] and Jingjing Ren[1]

Abstract

Objective: The primary aim of this study is to analyze health information seeking behaviors of users related to child fever within online health communities. The findings will serve as a foundation for the development of targeted interventions and resources for addressing the specific information needs related to child fever. Ultimately, this will enhance parental capabilities in managing fever in children and for improving the quality of communication between healthcare professionals and parents dealing with feverish children.

Methods: This study employed data crawling to gather Q&A data on childhood fever from online health communities, specifically "haodf.com" between March 15, 2022, and March 15, 2023. A total of 47,781 texts were analyzed using a mixed research approach that combines qualitative text topic analysis with BERTopic algorithm.

Results: The health information needs regarding children's fever can be categorized into 6 primary topics and 17 secondary topics. Among them, parents' demand for medication consultation and medical guidance (Topic A) was the highest at 45.40%, followed by information concerning the management of fever symptoms and body temperature in children (Topic B) at 30.35%. 13.24% of the data focused on examination recommendations and interpretation of results (Topic C).

Conclusions: This study proposes a mixed thematic analysis method combining qualitative text thematic analysis and the BERTopic topic model, which reveals parents' information-seeking behaviors about children with fever. It emphasizes the challenges faced by parents in assessing their children's condition and highlights the necessity of continuous health information support and evidence-based medical knowledge. This can promote the improvement of medical services, optimize doctor–patient communication, strengthen patient information support, and optimize the content of online health communities.

Keywords

Health information seeking behaviors, childhood fever, online health community, topic analysis, BERTopic

Submission date: 23 March 2024; Acceptance date: 16 August 2024

Introduction

Fever is one of the most common symptoms in pediatric outpatient and emergency clinics. In China, the annual incidence of febrile illnesses among children exceeds 60%, and approximately half of all children experience fever one to five times per year. In the United States, approximately 60 million children visit the doctor due to fever each year.[1–3] Although febrile illnesses are typically benign and self-limiting, they remain the leading cause of death

[1]The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[2]Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China
[3]School of Information Technology, Zhejiang Financial College, Hangzhou, China

Corresponding author:
Jingjing Ren, The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou 310003, China.
Email: 3204092@zju.edu.cn

for children under five globally.[4,5] However, numerous studies have shown that parents' knowledge and practical skills regarding childhood fever and its management are still very limited. They often display excessive anxiety and fear when faced with their child's fever, leading to unnecessary emergency visits and repeated medical consultations.[6–8] On the other hand, there are still children who die from severe infections due to delayed medical attention.[9]

Parents can obtain information on managing childhood fever through various channels such as medical staff, relatives and friends, books, television, and the internet. Generally they indicate that information from medical staff is relatively authoritative and trustworthy, but the content is often too specialized, which makes it difficult for them to understand and apply in practice. Additionally, inconsistencies in information from different sources significantly exacerbate parents' confusion and anxiety about fever management.[2,10–12] Greenhalgh[13] discussed the importance of including patients in quality improvement plans in practice when implementing evidence-based healthcare. Evidence-based healthcare requires identifying patients' needs and what is important to them. Therefore, conducting research to assess parents' information needs for childhood fever management to ensure evidence-based information support is crucial for safe and effective fever management by parents. Most previous studies relied on traditional interventional surveys based on questionnaires or interviews. These studies often had limited sample sizes, and the information gleaned from these designed questionnaires and interviews tends to be constrained, and the potential for researcher bias is significant, thereby limiting the scope of the findings and introducing possibilities for concealed biases.[3,7,11,14]

With the deepening development of "Internet + healthcare," the public is increasingly proactive in obtaining information and participating in medical and health decision-making through online channels. Meanwhile, health information users leave massive and rich data after using information on the platform, providing the most direct and convenient data source for analyzing users' health information needs and behaviors.[15] Text topic analysis is widely used in the content analysis of textual materials. Early text topic analysis focused on qualitative research methodologies, wherein researchers manually encoded texts to statistically analyze and extract key semantic features. The encoding process can be broadly categorized into two distinct approaches: deductive and inductive coding. In deductive coding, researchers construct categories based on prior knowledge or experience and then encode the text based on these predefined classifications. Conversely, inductive coding requires researchers to initiate the coding process from the ground up, without relying on any preexisting category system. In this approach, category formation and text encoding occur simultaneously. Most qualitative research refers to an inductive coding process rather than a deductive one. Although qualitative topic analysis offers the advantage of more accurately extracting thematic meanings from textual data, it demands a significant investment of human resources and time.[16,17]

The vast amount of information within online health communities makes it challenging to conduct text thematic research solely through manual coding. Text mining techniques utilizing computerized automation methods are more suitable for analyzing large volumes of online textual data.[18] BERTopic[19] stands as an advanced computerized topic modeling algorithm. Compared to the traditional Latent Dirichlet Allocation topic model,[20] it incorporates the BERT pre-trained deep learning model, exhibiting stronger semantic comprehension. It understands the contextual meaning of words, making it particularly suitable for topic mining in short texts, such as social media tweets and comments. Additionally, BERTopic can automatically determine the number of topics without presetting, thereby better assisting researchers in identifying optimal results. Despite its high efficiency in topic model analysis, BERTopic has certain limitations. Both the determination of a custom vocabulary, adjustment of model parameters, further induction and fusion of topics, and the naming of final topics require manual analysis from researchers. Therefore, this study adopts a mixed methodology combining qualitative text thematic analysis and BERTopic modeling. By integrating manually established topic categories with the experimental results from BERTopic, we aim to leverage the strengths of both approaches, enabling a more effective analysis of Q&A data themes. This approach provides a more comprehensive and authentic representation of health information seeking behaviors of users related to children's fever.

Methods

The overall research process is segmented into four steps (Figure 1): data collection, data preprocessing, topic extraction, and results discussion. Within the discussion of the results, two subsections are included: statistical analysis and topic analysis. The statistical analysis section uses descriptive statistics to provide an overview of users in online health communities who are seeking information on childhood fever. It also identifies potential factors that may influence their information-seeking behavior. The topic analysis section explores the characteristics of information needs related to childhood fever, highlighting the distribution and trend features of these needs.

Data collection

This study initially referred to the comprehensive rankings of medical and health industry websites on a Chinese website ranking platform, which combines Alexa rankings, Baidu weight, PR values, and mobile usage results. We manually screened the top 20 online health communities,
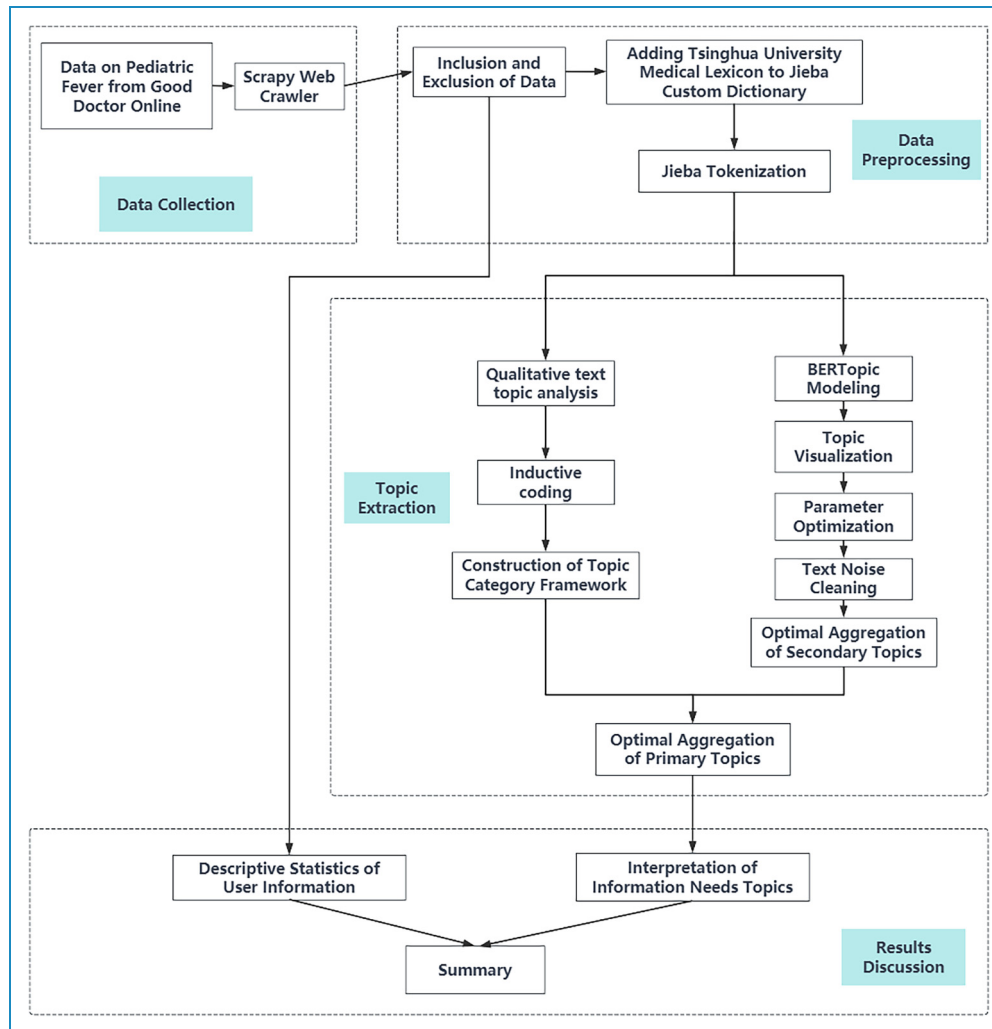
Figure 1. Research objectives flow.

including "haodf.com, XYWY, and Dxy.com." Ultimately, "haodf.com" was chosen as the data source. We utilized the Scrapy framework to construct a web crawler tool, which scraped 47,781 doctor–patient Q&A texts related to the topic of childhood fever from "haodf.com" between March 15, 2022, and March 15, 2023. BeautifulSoup was employed for web page parsing, extracting attributes such as user gender, age, height, weight, disease description, duration of illness, medical consultation status, medication status, medical history, allergy history, assistance sought, doctor's advice, basic information about the responding doctor, and response time. All this information was stored in a MongoDB database for subsequent processing and analysis.

Data preprocessing

Data inclusion and exclusion. Initially, this study selected data from pediatric patients aged between 0 and 5 years.

Notably, in December 2022, influenced by the relaxation of domestic COVID-19 restrictions, there was a surge in online consultations, with a single-month consultation count reaching 8319, accounting for 40.50% of the annual consultation volume (20,556 cases). To ensure data balance and to prevent the concentration of data from a specific period affecting the analysis results, we decided to extract 1112 cases from the 8319 cases in December using a random number table method, based on the average monthly consultation volume for the year (1112 cases). Additionally, through manual screening, low-quality data that were irrelevant to the research topic or incomplete were excluded to ensure data quality and research accuracy. Ultimately, a total of 12,583 cases were included in the analysis for this study.

Adding Tsinghua University Medical Lexicon to Jieba's custom dictionary. To enhance the accuracy of word segmentation, we incorporated the Tsinghua University Medical Lexicon

into Jieba's custom dictionary. This allows Jieba to better recognize medical terminology, thereby improving the precision of word segmentation.

Jieba segmentation. We used Jieba for segmentation processing, which is a Python-based Chinese segmentation tool that can break a continuous text into a series of individual words.

Topic extraction

Qualitative text topic analysis. According to Braun's six-step thematic analysis method,[21] which includes familiarizing yourself with your data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report, we analyzed the data. The coding task was collaboratively accomplished by two female professionals, both specializing in disease coding. Firstly, the researcher and another medical master carefully browsed and read the entire dataset. Then, 50 data samples were randomly selected for pre-coding. The two coders performed coding separately, and Cohen's kappa coefficient was calculated as a reliability indicator between the two coders. Using IBM SPSS 23.0 software to calculate the pre-coding results, the Cohen's kappa coefficient was 0.878, which is greater than 0.80. This proves that the two coders have a high degree of consistency in understanding the data, making it possible to proceed with formal coding.[22]

The two coders performed inductive coding on the text data separately by Microsoft 365 Excel, and then compared the initial coding results. In case of disagreement, they negotiated to reach a consensus on the coding results. When no new semantic expressions appeared, the data was considered saturated, and coding was stopped. A total of 420 data samples were coded.[23] A total of 618 health-related questions have been encoded. Codes with similar or closely related meanings were further summarized to form 10 secondary topics and 6 primary topics. This study referenced a variety of clinical guidelines, expert consensuses, and research literature related to childhood fever. These included the expert consensus on the rational use of antipyretic analgesics in symptomatic treatment of childhood fever,[24] 30 questions on childhood fever health education,[25] and guidelines for diagnosing and treating acute fever of unknown etiology in children aged 0 to 5 in China (standard version).[26] We also referred to the expert consensus on pediatric emergency pre-examination and triage in Shanghai[27] and the treatment of high fever in children.[28] Based on the actual questions posed by parents, the content was categorized into 6 primary topics and 10 secondary topics (Table 1).

Topic clustering via BERTopic. The algorithmic execution flow of BERTopic is illustrated in Figure 2.[29]

In this study, we employed the "Paraphrase-multilingual-MiniLM-L12-v2" word vector model to generate embedded representations of the preprocessed text. During the processing of the textual features, we utilized the CountVectorizer function. This approach enabled us to compute the frequency of each topic within individual documents and to impose restrictions on the term frequency within the documents.

We relied on the similarity matrix derived from the BERTopic model, without a predetermined number of clustering topics. By integrating the UMAP and HDBSCAN algorithms, topics with higher similarity were merged. It's crucial to note that the HDBSCAN algorithm might produce outliers, which implies that certain documents might not be associated with any specific topic. For these outlier documents, we employed BERTopic's "reduce_outliers" function to calculate their likelihood of belonging to each topic and subsequently assigned them to the topic with the highest probability.

In the phase of determining the number of topics, we initially refrained from adjusting any parameters, allowing the model to autonomously generate relevant topics. This procedure culminated in the identification of 229 topics, which were interpreted as tertiary topic outcomes, as depicted in Figure 3(a). Upon observing the topic map, it's evident that a significant overlap exists among most topics, which suggests the potential for enhanced clustering through further aggregation. Drawing insights from qualitative text topic analysis results, we embarked on further topic optimization and aggregation. By utilizing BERTopic's "reduce_topics" function, we ascertained that when the topic count is set at 17, the outcomes can be classified as secondary topics. The distribution of these secondary topics is illustrated in Figure 3(b).

However, we observed from the topic distribution map (Figure 3(b)) that some topics are very close to each other, suggesting that there is still room for further consolidation to extract primary topics. Therefore, by combining the results of qualitative text topic analysis categorization and referring to the topic similarity matrix (Figure 4), and through expert discussions, we ultimately decided to use the categorization method provided in BERTopic.

To better understand the interrelationships between topics, we adopted a visualization-based topic reduction approach. From the secondary topic similarity matrix (Figure 4), we can see that Topic 0 (Medication Guidance and Medical Indications) and Topic 6 (Application of Antibiotics) have a high similarity of 0.71. From a professional knowledge perspective, the use of antibiotic drugs also falls under the broader category of medication use, and they can be merged into a single topic. Similarly, Topic 4 (Blood Tests and Results) and Topic 9 (White Blood Cell Count and Analysis) can also be combined into one topic. After sequentially reviewing the structure of these topics, we finally determined 7 primary topics. The integration results of qualitative text topic analysis and BERTopic topic modeling are as follows (Figure 5).

Table 1. The categorization strategy and topic system for information needs on childhood fever ($N = 618$).

Primary topics	Secondary topics	Categorization strategy
Etiology ($n = 27$, 4.37%)	Etiology	Etiologies leading to fever in children, such as vaccination, exposure to cold, dietary factors, etc
Diagnosis ($n = 112$, 18.12%)	Disease diagnosis ($n = 72$, 11.65%)	Diagnosis of primary diseases causing fever
	Auxiliary examination ($n = 40$, 6.47%)	Recommendations for auxiliary examinations and interpretation of their results
Treatment ($n = 457$, 73.95%)	Medication guidance ($n = 208$, 33.66%)	Inquiry into the types of antipyretics and medications for treating the primary disease, including dosage, frequency, combination therapies, methods of administration, observation and evaluation of drug efficacy, and adverse reactions to medications
	Medical indications ($n = 94$, 15.21%)	Considerations including outpatient visits, indications for hospitalization, and referrals
	Symptomatic treatment ($n = 83$, 13.43%)	Clinical manifestations and management of fever and related symptoms
	Treatment plan ($n = 72$, 11.65%)	Comprehensive therapeutic recommendations based on the child's disease condition
Complications and comorbidities ($n = 12$, 1.94%)	Complications and comorbidities	Assessment of the risk of complications related to fever, as well as the diagnosis and treatment of comorbidities and complications
Nursing and rehabilitation ($n = 8$, 1.29%)	Nursing and rehabilitation	Precautions for home care of febrile children
Prevention ($n = 2$, 0.32%)	Prevention	Inquiry into the child's exposure to hazardous environments, such as parents with fever, or recurrent fevers, and measures to prevent fever or illness in children

The distribution map of these primary topics, as illustrated in Figure 5(b), is relatively dispersed, displaying a desirable clustering effect. Throughout this process, we maintained flexible control over the number of topics and continuously optimized the clustering results. This approach not only improved the clustering effect of the topics but also clarified the hierarchy of the topics, aiding in our better understanding and interpretation of the model's outcomes.

Each topic is represented by multiple terms, with each term being represented by a c-TF-IDF score. The higher the score, the greater the contribution of the term to the topic. The terms for each topic are sorted based on their c-TF-IDF scores, which decrease with each additional term. Figure 6 illustrates the decline of term scores with c-TF-IDF for each topic. From this, we can observe that the first three keywords of each secondary topic can

essentially represent the main content of the topic. Based on this, naming was assigned to each primary and secondary topic, with the results and the main topic keywords presented in Table 2.

Results

Excluding 15 cases of noisy interfering data, a total of 12,568 cases of data were included in this study for topic analysis.

Descriptive statistics of general information queried from users

According to the statistical results, the number of questions about male and female pediatric patients were 6914 (55.01%) and 5654 (44.99%), respectively, resulting

in a gender ratio of 1.22:1, indicating a slight predominance of males (Figure 7(a)). The majority of the children, accounting for 88.46% (11,118), were aged 0–3 years (Table 3). Among them, the highest proportion of consultations came from 1-year-old toddlers, accounting for 29.66%. The age-specific consultation volume exhibited a unimodal characteristic (Figure 8), with a declining trend in consultation volume after the age of 2. Among the children who sought advice in the community, 4514 (39.92%) confirmed that they had previously visited a hospital, 181 (1.44%) confirmed they had never visited, and 7873 (62.64%) did not mention their medical visit status (Figure 7(b)).
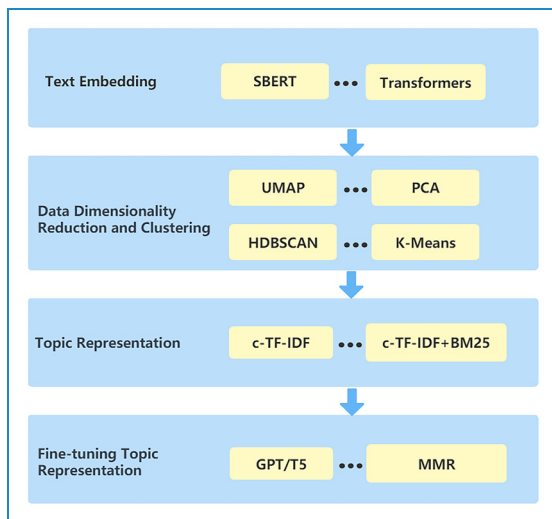

Figure 2. BERTopic workflow.

Topic analysis results

The research identified that the health information needs related to childhood fever can be divided into 6 primary topics and 17 secondary topics (Table 4). Out of the 12,568 coded data entries, 5706 cases (45.40%) were related to Topic A: Medication Consultation and Medical Guidance. The keyword cloud for Topic A (Figure 9), includes drug selection, the application of antibiotics, indications for intravenous fluids, and indications for outpatient, inpatient, and referral medical visits. Therefore, guidance on drug treatment and medical decision-making emerges as the most critical information needs for parents of febrile children. Additionally, 3815 cases (representing 30.35% of the sample) focused on Topic B: Symptoms and Temperature Management of Febrile Children, ranking it second. This topic covers symptomatic treatment for fever and recurrent fever, the application of antipyretics, changes in nighttime conditions, and the management of worsening fever symptoms. Topic C, with 1664 cases (13.24%), pertains to Examination Suggestions and Result Interpretation. According to the Pareto Principle (80/20 rule), factors that accumulate within the 0–80% range are considered primary factors.[30] Topics A, B, and C together account for 88.99% of the total, making them the most significant information needs. In Topic D: Etiological Diagnosis and Treatment, parents mainly inquire about the definitive diagnosis of diseases causing fever symptoms, such as pediatric emergencies, rashes, H1N1, COVID-19, colds, urinary tract infections, and non-disease factors like vaccine-induced fever, especially measles-mumps-rubella vaccine, allergies, cold exposure, and overeating. Diet and feeding of infants and young children are the most concerning content for parents in the topic
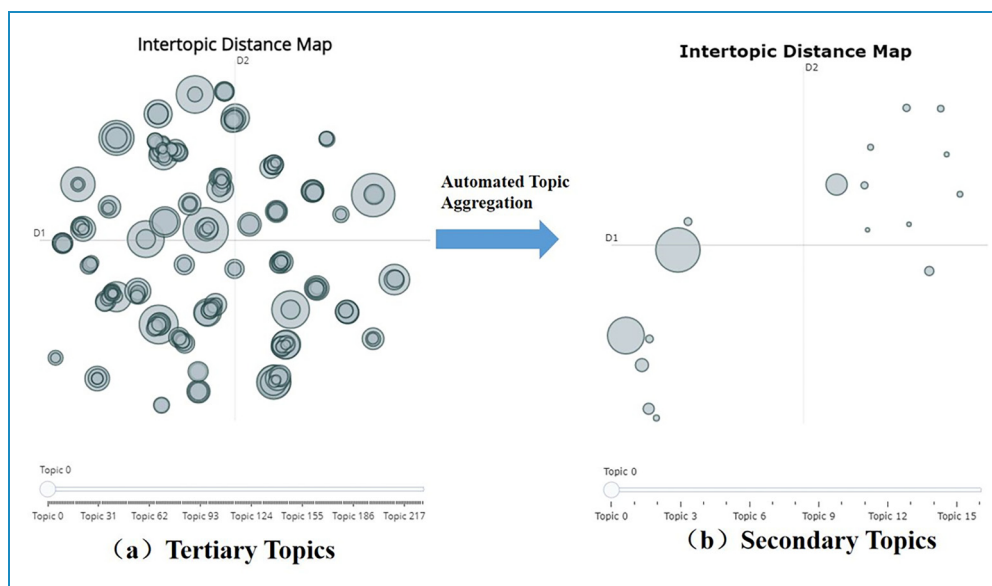

Figure 3. Evolution of the distribution of secondary topics.
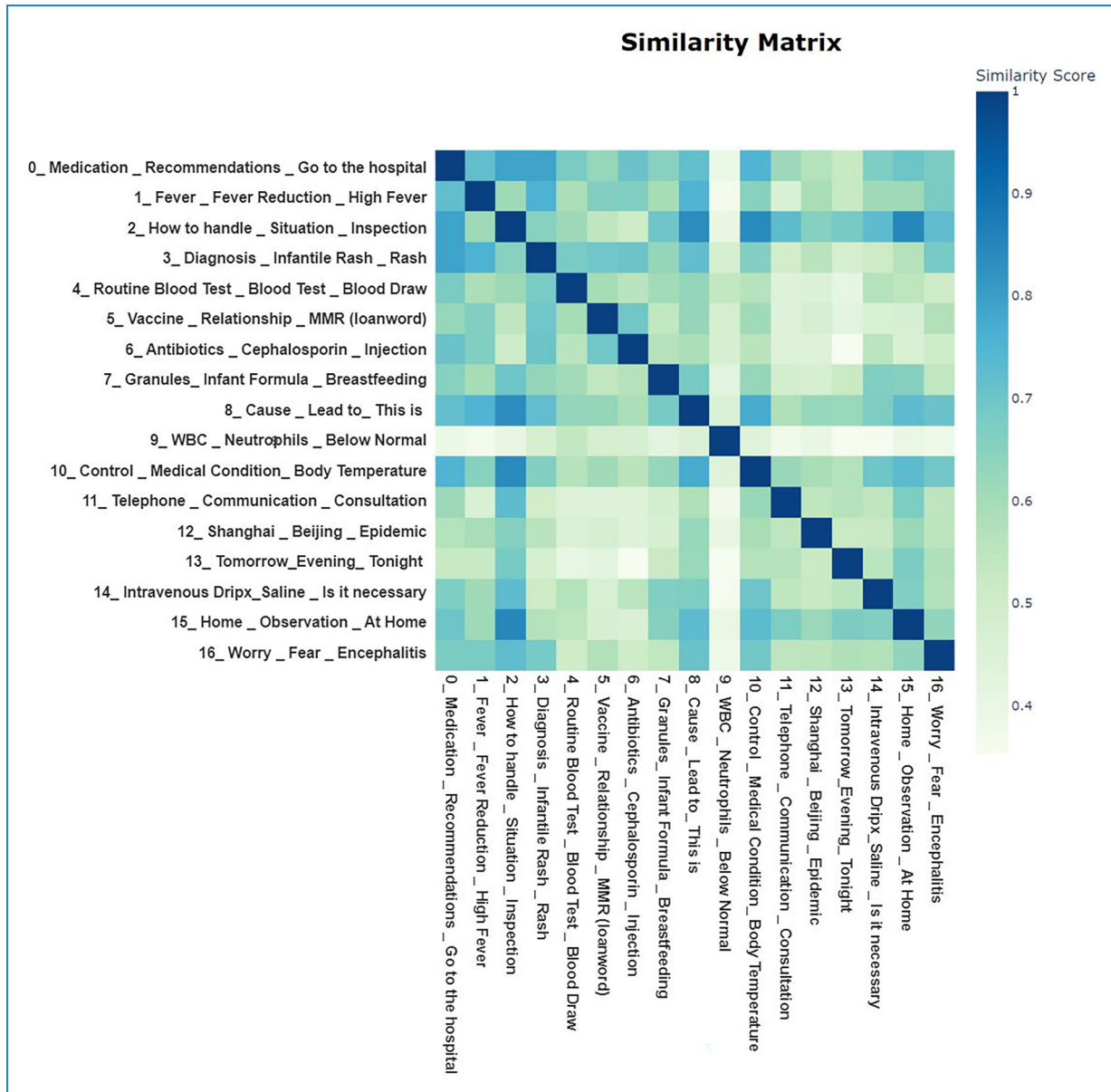
Figure 4. Secondary topics similarity matrix.

of Rehabilitation Management and Home Care, especially breastfeeding issues during the child's fever or medication period. Additionally, home observation of the child's condition, the use of antipyretics, and nebulization for phlegm removal and cough suppression are also the main focuses of parents in this topic. 86 cases (0.70%) of parental inquiries were identified as having negative emotions such as worry, fear, and anxiety, mainly related to concerns about severe complications of fever, such as encephalitis, convulsions, and myocardial injury. The last topic, Topic F: Telephone Communication and Consultation, with 53 cases (0.42%), involves parents hoping to communicate directly with doctors by phone. On the one hand, due to the urgency of the situation, they hope to get a quick response;

on the other hand, they are concerned that textual descriptions may not be exhaustive and prefer direct communication with the doctor.

Discussion

Qualitative text topic analysis and topic clustering via BERTopic: synergistic dual tracks for advanced textual exploration

By contrasting the proportion of thematic distributions between the two methodologies, it was observed that the qualitative research approach yielded the highest prevalence
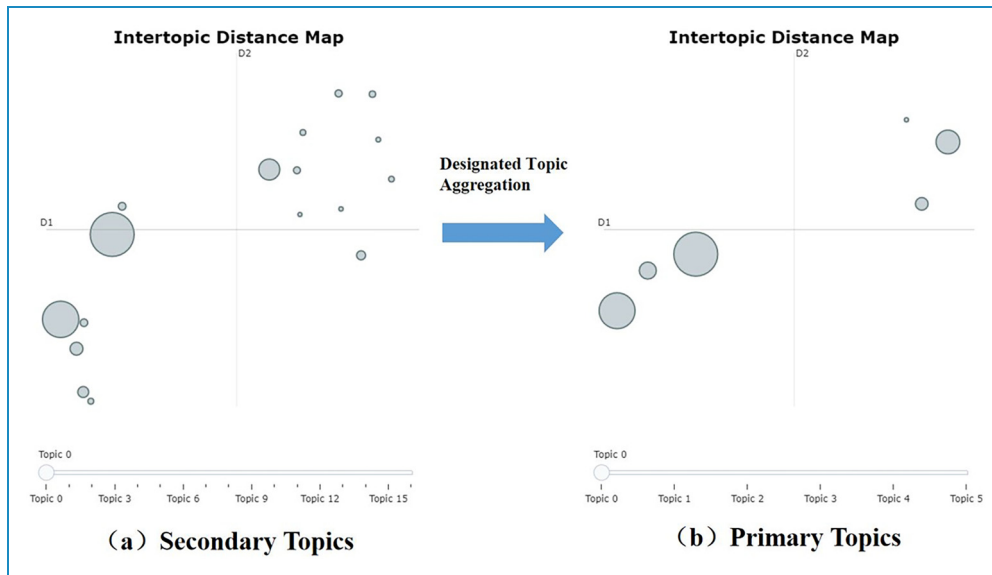
Figure 5. Evolution of the distribution of primary topics.
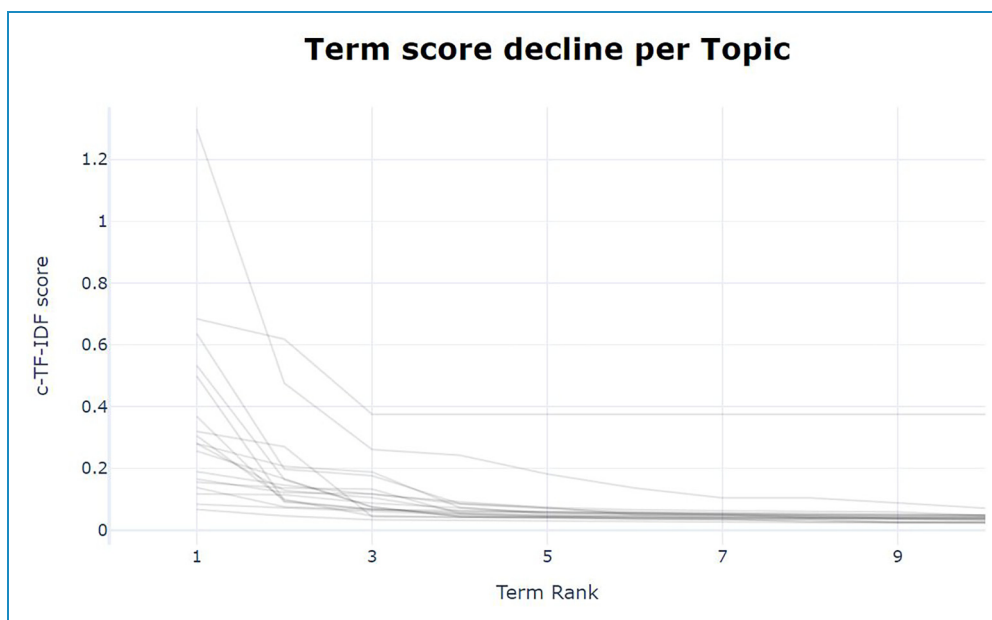
Figure 6. Term score decline per topic.

of medication guidance themes, followed by medical advice seeking and symptomatic management, which aligns closely with the top three themes unearthed by the BERT model. From a thematic content perspective, the qualitative method meticulously segregates and clarifies each inquiring parent's diverse needs individually, whereas the BERTopic algorithm emphasizes an overall data distribution, consolidating multiple needs from a single patient. This results in a reduced total count of primary themes but encompasses a broader scope of content that aligns more closely with real-world scenarios. Regarding the delineation of secondary themes, the BERTopic algorithm, leveraging big data, exhibits a broader thematic coverage and a more even data distribution in its classification outcomes. The themes uncovered, such as "Management of Nocturnal Fever," "Disease Diagnosis and Pediatric Rash," "Disease Anxiety and Disease Management," and "Telephone Communication and Consultation," are not only detailed but also enriched, offering additional thematic content and nuances that enhance their practical application value.

In summary, the findings from qualitative research provide a solid foundation for the application of the

Table 2. Topic keywords data output from the BERTopic model.

Primary topics	Secondary topics	Topic keywords
Medication Consultation and Medical Guidance	0. Medication Guidance and Indications for Medical Consultation	Medication, Recommendation, Hospitalization, Treatment, Medical Consultation, Medical Condition, Which Medication to Take, Control, Examination, What Medication
	6. Application of Antibiotics	Antibiotics, Cephalosporins, Injection, Azithromycin, Bacterial Infection, Infusion, Oral Administration, Intake, Hospital, Cefuroxime
	12. Urban Epidemic Medical Consultation and Referral	Shanghai, Beijing, Epidemic, Hospital Transfer, Children's Hospital, Anti-inflammatory, Hospitalization, Medical Visit, Hospital, Examination
	14. Indications for Intravenous Infusion and Inpatient Treatment	IV Drip, Saline, Is it Necessary, IV Fluids, Ineffective, Is it Suitable, Hospitalization, Medication, Viral, Not Improving
Fever and Temperature Management	1. Fever Symptoms and Methods of Fever Reduction	Fever, Fever Reduction, High Fever, Pyrexia, Body Temperature, Recurrent, Cause, Going to the Hospital, Cough, Recurrent Fever
	13. Management of Nocturnal Fever	Tomorrow, Evening, Tonight, Midnight, Morning, Not Well, Daytime, Corticosteroids, Method, Night
Inspection Recommendations and Result Interpretation	2. Medical and Inspection Guidance	How to Handle, Situation, Examination, Recommendation, Further, Meilin (a medication), Guidance, Opinion, Take a Look, Oseltamivir
	4. Blood Tests and Results	Complete Blood Count, Blood Test, Blood Draw, Blood, Laboratory Test, Myocardium, Blood Parameters, Report, Take a Look, Examination
	9. White Blood Cell Count and Analysis	White Blood Cells, Neutrophils, Below Normal, Granulocytes, Platelets, Above Normal, Count, Monocytes, Cause, Cephalosporins
Etiological Diagnosis and Treatment	3. Disease Diagnosis and Pediatric Rash	Diagnosis, Roseola Infantum, Rash, Etiology, Allergy, Disease, Medical Condition, Confirmed Diagnosis, Treatment, Judgment
	5. Identification and Management of Vaccine-induced Fever	Vaccine, Relationship, Mumps-Measles-Rubella Vaccine, Immunity, Cause, COVID-19, Fever, Is it Due to, Common Cold, Vaccination
	8. Etiological Diagnosis of Fever	Cause, Leading to, This is, Fever, Urinary Tract Infection, Pyrexia, High Fever, Analysis, Treatment, Recurrent
Rehabilitation Management and Home Care	7. Infant and Toddler Nutrition and Feeding	Granules, Infant Formula, Breastfeeding, Breast Milk, Chinese Traditional Medication, Some Food, Breast Milk, Traditional Chinese Medicine Granules, Drinking Milk, Impact
	10. Disease Control and Temperature Management	Control, Medical Condition, Body Temperature, Inflammation, Etiology, Streptococcus, Examination, Reduce, Recommendation, Better
	15. Home Observation and Care	At Home, Observation, At Home, Nebulization, Tylenol, Going to the Hospital, Going Home, Hospital, H1N1 Influenza, Oseltamivir

Table 2. Continued.

| Primary topics	Secondary topics	Topic keywords
	16.Disease Anxiety and Disease Management	Worry, Fear, Encephalitis, Worsening, Probability, Convulsion, Anxiety, Disease, Method, Use Some
Telephone communication and consultation	11.Telephone Communication and Consultation	Telephone, Communication, Consultation, Cefaclor, Details, Interaction, What Time, Nine O'clock, Pediatrician, Guidance
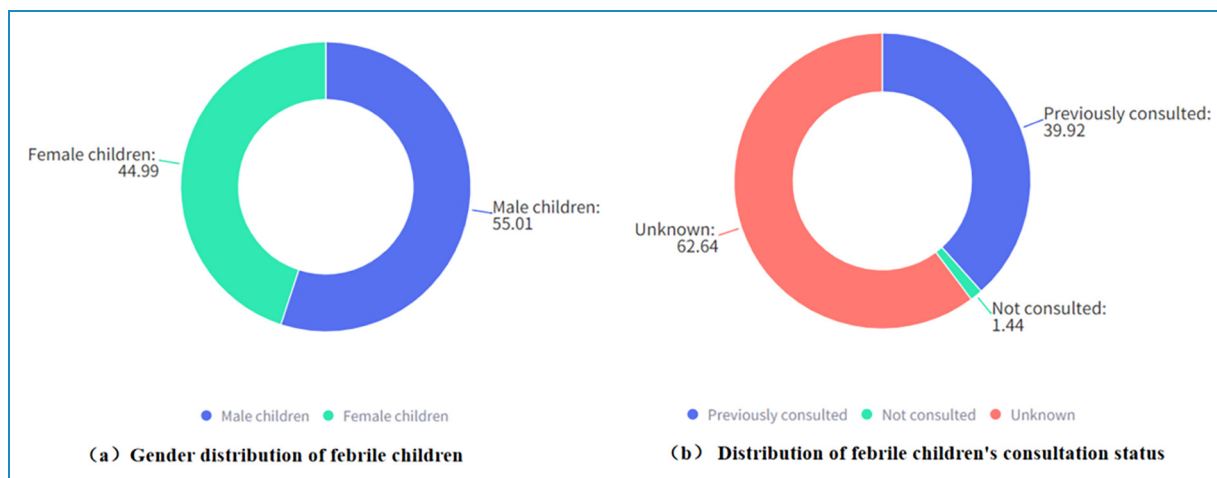
Figure 7. Distribution of gender and consultation status among febrile children.

Table 3. Age distribution of febrile children in online community consultations.

Age group (years)	Number of consultations (cases)	Proportion	Cumulative proportion
0–1	2810	22.36%	22.36%
1	3728	29.66%	52.02%
2	2518	20.04%	72.06%
3	2062	16.41%	88.46%
4	1352	10.76%	99.22%
5	98	0.78%	100.00%

BERTopic algorithm. By juxtaposing the outcomes of both, researchers can fine-tune and optimize the BERTopic algorithm, ensuring its outputs better mirror actual professional requirements and user expectations. This not only enhances the applicability and reliability of the BERTopic algorithm but also strengthens its capacity to comprehend specialized terminology and subtle distinctions, thereby more accurately reflecting users' information needs and preferences in real-world applications. This synergy of artificial intelligence and human expert knowledge underscores the formidable potential of modern technologies in healthcare research and practice.

Families of febrile children need ongoing health information support

Children with febrile conditions often exhibit rapid progression and recurrent symptoms. Throughout the treatment and recovery process, parents require extensive health information to aid them in decision-making and symptom management. In this study, out of 4514 cases (39.92%) that confirmed previous medical consultations, children still encountered health issues requiring medical consultation during their home recovery. Evidently, from the primary and secondary topic content, parents' health information needs span the entire phase from disease diagnosis, treatment, to recovery, and these needs dynamically change at different stages. This result is consistent with two qualitative research findings reported abroad,[31,32] despite the lack of knowledge about fever being the most important reason leading to anxiety and fear, most parents initially have confidence in caring for children with fever. However, changes in the condition of the fever, especially the
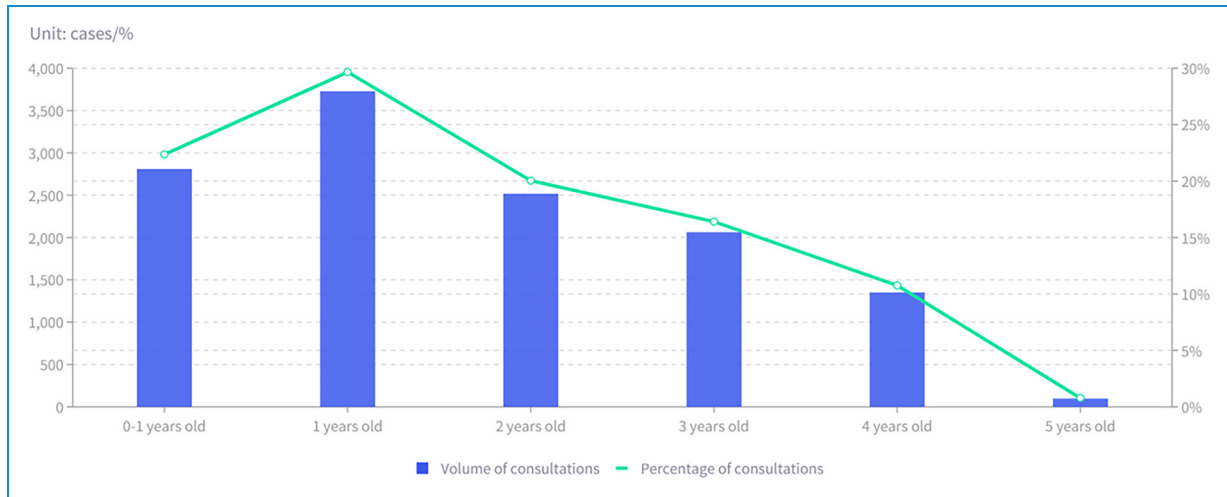
Figure 8. Age distribution of febrile children in online community consultations.

appearance of new symptoms and signs, trigger parents' unease, causing them to question their knowledge and skills and exacerbating panic emotions. The changes and development of the febrile condition increase parents' need for information support and also become a turning point for parents to seek medical healthcare services for their children. However, the reality is that there is a limitation in pediatric medical resources in China. Common challenges include difficulty in making appointments, long waiting times, extended periods for test appointments and result releases, and short consultation duration. Relying solely on hospital visits is hardly sufficient to meet the needs of families throughout the illness journey.[33] Hence, establishing high-quality, timely, and accessible health information channels for these families is crucial to address this supply-demand imbalance. Currently, internet hospitals and online health communities are the mainstream professional channels providing health information to these families. Especially during the COVID-19 pandemic, these platforms effectively reduced the risk of cross-infection for children and parents visiting hospitals, alleviating their fears and anxieties.[34] However, internet hospitals and online health communities, which are in the developmental stage, still face challenges such as inconsistent information quality, low precision in information search, slow doctor response rates, and low patient satisfaction.[35,36] Therefore, there is an urgent need to construct a user-oriented and targeted "Internet Plus" medical service to provide patients with more convenient and high-quality health information support.

Parents generally lack the ability to assess the condition of their sick children

From the word cloud of drug consultations and medical guidance in Topic A, it is observed that "go to the hospital" is the third most frequent keyword, following "medication"

and "proposal." This phenomenon reflects the common confusion among parents about when to take their child to the hospital. The primary reason for this uncertainty is often the parents' lack of adequate knowledge about pediatric fever, which prevents them from accurately assessing the severity of their child's illness.[37] Previous surveys have shown that, compared to attending physicians, parents overestimate the severity of their child's condition by 45.9%, leading to 44.62% of children seeking medical attention prematurely.[38,39] On the other hand, the COVID-19 pandemic has also influenced parents' decisions about seeking medical care. Parents are concerned about potential cross-infections when taking their child to the hospital and the inconveniences caused by quarantine measures. This has made parents more rational and cautious when deciding whether to take their child to the hospital.[40]

Therefore, in providing education and guidance to parents of febrile children, medical professionals should emphasize the key points of home-based symptom observation and the indications for seeking medical attention. This approach aims to prevent children from seeking medical care too early, delaying treatment, or repeatedly visiting the hospital. By doing so, not only can parents better understand and manage their child's condition, but it can also effectively alleviate the strain on the healthcare system, leading to a more efficient allocation of medical resources.

Parental decision-making needs to be empowered by evidence-based medical knowledge

When it comes to the management of fever in toddlers, parents' decisions and interventions are extremely cautious. From Topic B, it is evident that parents have a significant need for information on how to manage fever symptoms in their children. Although parents can easily obtain relevant information from family, friends, and the internet,

Table 4. Topic information mined based on qualitative text topic analysis and BERTopic algorithm ($N = 12{,}568$).

Primary topics and secondary topics	Quantity n (%)
A. Medication Consultation and Medical Guidance ($n = 5706$, 45.40%)	
0.Medication Guidance and Indications for Medical Consultation	5431 (43.21)
6.Application of Antibiotics	165 (1.31)
12.Urban Epidemic Medical Consultation and Referral	62 (0.49)
14.Indications for Intravenous Infusion and Inpatient Treatment	48 (0.38)
B. Fever and Temperature Management ($n = 3815$, 30.35%)	
1.Fever Symptoms and Methods of Fever Reduction	3723 (29.62)
13.Management of Nocturnal Fever	92 (0.73)
C. Inspection Recommendations and Result Interpretation ($n = 1664$, 13.24%)	
2.Medical and Inspection Guidance	1233 (9.81)
4.Blood Tests and Results	337 (2.68)
9.White Blood Cell Count and Analysis	94 (0.75)
D. Etiological Diagnosis and Treatment ($n = 859$, 6.83%)	
3.Disease Diagnosis and Pediatric Rash	471 (3.75)
5.Identification and Management of Vaccine-induced Fever	160 (1.27)
8.Etiological Diagnosis of Fever	228 (1.81)
E. Rehabilitation Management and Home Care ($n = 471$, 3.75%)	
7.Infant and Toddler Nutrition and Feeding	141 (1.12)
10.Disease Control and Temperature Management	111 (0.88)
15.Home Observation and Care	133 (1.06)
16.Disease Anxiety and Disease Management	86 (0.70)
F. Telephone communication and consultation ($n = 53$, 0.42%)	
11.Telephone Communication and Consultation	53 (0.42)

the reliability of this information is not always guaranteed. Parents tend to seek advice from medical professionals, especially regarding medication use.[32] With the rise in public health awareness, parents increasingly wish to be involved in health care decisions and activities for their children, choosing the most appropriate diagnostic and treatment options. In this study, 1664 parents (representing 13.24% of the sample) sought advice on examinations and desired a more in-depth and comprehensive analysis of the examination reports they obtained. This is to avoid excessive examinations that might harm the child and further confirm the accuracy of the current disease diagnosis and treatment plan. Research has shown that parents who are highly involved in decision-making can better understand the condition and treatment plan, thereby improving treatment adherence and outcomes. Their proactive attitude and trust in the treatment also help in timely treatment, avoiding additional medical expenses, effectively reducing

Figure 9. Topic A keyword word cloud.

medical costs.[41] The foundation for parents' participation in medical decisions is their possession of some evidence-based medical knowledge related to the disease, which allows them to balance evidence with personal wishes. Medical staff guidance and fever management guidelines are reliable sources of evidence-based medical knowledge for parents. However, this guidance, being highly professional and lacking a practical perspective, often becomes somewhat detached from the actual needs of patients, making it difficult for them to understand and apply in practice. For example, a parent's consultation stated, "I gave Tylenol two days ago, and Meilin last night. I don't know if the medication or the frequency was wrong, causing the fever to recur." In the "Expert Consensus on the Rational Use of Antipyretic Analgesics in the Symptomatic Treatment of Fever in Children,"[24] there is only a recommendation for the maximum dosage of antipyretics per dose and per day, lacking practical guidance on observing the effects of the medication and further management. Therefore, transforming evidence-based knowledge related to the diagnosis and treatment of febrile children into a format that is easy for parents to understand and apply is the key to empowering parents to participate in children's medical decisions.

Conclusions

In this study, building upon existing topic mining methods, we propose a novel text topic mining strategy that integrates the strengths of qualitative text topic analysis and BERTopic topic modeling. Qualitative text topic analysis,

carried out by medical professionals, allows for a deep dive into the text, parsing context and intricate semantics. On the other hand, BERTopic topic modeling offers a broad understanding as it can handle vast amounts of data, revealing topics and trends. Using a mixed-methods approach achieves a fusion of depth and breadth, balancing accuracy with efficiency. They validate and complement each other, offering greater flexibility and enhancing the interpretability of results. This approach capitalizes on the advantages of both, elevating the quality and efficiency of the analysis. Identifying patient needs is a crucial aspect of advancing healthcare service reforms. This study concentrates on the health information requirements of parents of children with fever in real-world settings. The insights gained may lay the groundwork for developing strategies aimed at enhancing doctor–patient communication or improving parents' ability to manage their child's fever. Furthermore, by scientifically categorizing and labeling unstructured text within health communities, our approach facilitates the statistical analysis and content optimization of web data. This refined categorization will enhance the website's search functionality, enabling patients to more precisely locate the information they require.

Limitations and future research

Due to the immense amount of information in the data source, the data collected in this study is solely sourced from the "haodf.com" online platform, and thus may not comprehensively reflect the health information needs of

parents regarding children's fever across all online health communities. Considering that the period of this study is one year, combined with the surge in data volume due to adjustments in China's COVID-19 control measures, we adopted a random sampling strategy for data selection. As a result, this study did not conduct a seasonal analysis, potentially failing to reveal health information demand trends over extended periods or seasonal variations. The findings indicate that many parents, when seeking health consultations, often exhibit emotional responses such as anxiety and panic. Based on this, future research should consider incorporating sentiment analysis techniques to further deepen the understanding of parents' health information needs and their areas of concern.

Consent to participate: "Not applicable," as all information utilized in this study was derived from publicly accessible online health communities, and none of the data pertained to personal privacy. Furthermore, the study did not entail any interaction with users.

Consent for publication: "Not applicable," as all information utilized in this study was derived from publicly accessible online health communities, and none of the data pertained to personal privacy. Furthermore, the study did not entail any interaction with users.

Contributorship: Xuan Pan did conceptualization, methodology, data curation, visualization, formal analysis, writing—original draft, writing—review and editing. Zhuoyuan Tang did conceptualization, methodology, data curation, visualization, writing—original draft, writing—review and editing. Ying Liu performed conceptualization, formal analysis, methodology, writing—original draft, writing—review and editing. Jingjing Ren performed conceptualization, formal analysis, funding acquisition, methodology, writing—original draft, writing—review and editing.

Data availability statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: No clinical data, human specimens, or laboratory animals were involved in this study. All information used in this study was obtained from publicly released online health communities, and none of the data involved personal privacy. In addition, the study did not involve any interaction with users; therefore, no ethics review was required.

Guarantor: Jingjing Ren.

ORCID iD: Xuan Pan ⓘ https://orcid.org/0000-0002-1306-5706

Supplemental material: Supplemental material for this article is available online.

References

1. Paul SP, Kini PK, Tibrewal SR, et al. NICE guideline review: fever in under 5s: assessment and initial management (NG143). *Arch Dis Child Educ Pract Ed* 2022; 107: 212–216.
2. Alsofyani BA and Hassanien NS. Factors affecting parent's practice regarding the management of children's fever. *Cureus* 2022; 14: e25658.
3. Liu XJ. Correlation study between parents' knowledge and cognition of fever, anxiety, and health-seeking behavior. Master Thesis, China: Yanbian University, 2022.
4. Bruno M, Ellis A, et al. Consenso para el uso adecuado de antibióticos en el niño menor de 36 meses con fiebre sin foco de infección evidente. Resumen ejecutivo [consensus for the proper use of antibiotics in children under 36 months with fever without source of infection. Executive summary]. *Arch Argent Pediatr* 2017; 115: 205–206.
5. Perin J, Mulick A, Yeung D, et al. Global, regional, and national causes of under-5 mortality in 2000–19: an updated systematic analysis with implications for the sustainable development goals. *Lancet Child Adolesc Health* 2022; 6: 106–115.
6. May M, Brousseau DC, Nelson DA, et al. Why parents seek care for acute illness in the clinic or the ED: the role of health literacy. *Acad Pediatr* 2018; 18: 289–296.
7. Daifallah A, Jabr R, Al-Tawil F, et al. An assessment of parents' knowledge and awareness regarding paracetamol use in children: a cross-sectional study from Palestine. *BMC Public Health* 2021; 21: 380.
8. Williams MR, Greene G, Naik G, et al. Antibiotic prescribing quality for children in primary care: an observational study. *Br J Gen Pract* 2018; 68: e90–e96.
9. Craig JC, Williams GJ, Jones M, et al. The accuracy of clinical symptoms and signs for the diagnosis of serious bacterial infection in young febrile children: prospective cohort study of 15 781 febrile illnesses. *Br Med J* 2010; 340: c1594.
10. Kelly M, McCarthy S, O'Sullivan R, et al. Drivers for inappropriate fever management in children: a systematic review. *Int J Clin Pharm* 2016; 38: 761–770.
11. Kerdar SH, Himbert C, Martin DD, et al. Cross-sectional study of parental knowledge, behaviour and anxiety in

management of paediatric fever among German parents. *BMJ Open* 2021; 11: e054742.

12. Walsh A, Edwards H and Fraser J. Over-the-counter medication use for childhood fever: a cross-sectional study of Australian parents. *J Paediatr Child Health* 2007; 43: 601–606.

13. Greenhalgh T. *How to implement evidence-based healthcare.* 1st ed. New Jersey: Wiley Blackwell, 2018, p.130.

14. Bobby A. *Social research methods.* 11th ed. Beijing: Huaxia Publishing House, 2000, p.318.

15. Zhou XY. Direction and focus of comprehensive management of online health information. *Governance* 2018; 47: 45–49.

16. Huang C and Lyu LY. The application of text analysis to public management and policy research. *Chin Public Adm Rev* 2020; 2: 156–175.

17. Fereday J and Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods* 2006; 5: 1–11.

18. Mu SY. Text theme and sentiment analysis of complaints on third-party complaint platforms. Master Thesis, China: Dongbei University of Finance and Economics, 2023.

19. Grootendorst M. BERTopic: neural topic modeling with a class-basedTF-IDF procedure, https://maartengr.github.io/BERTopic/ (2022, accessed 10 March 2023).

20. Blei DM, Ng AY, Jordan MI, et al. Latent dirichlet allocation. *J Mach Learn Res* 2003; 3: 993–1022.

21. Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3: 77–101.

22. Popping R. *Sociometric research.* 1st ed. London: Palgrave Macmillan, 1988, pp.90–105.

23. Hsieh HF and Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005; 15: 1277–1288.

24. China National Clinical Research Center for Respiratory Diseases, Group of Respirology, Chinese Pediatric Society, et al. Expert consensus on rational use of antipyretics and analgesics in the treatment of fever in children. *Chin J Appl Clin Pediatr* 2020; 35: 161–169.

25. China National Clinical Research Center for Respiratory Diseases, China Medicine Education Association Committee on Pediatrics, Chinese Medical Doctor Association Committee on Respirology Pediatrics, et al. Thirty key issues on management of febrile children. *Chin J Appl Clin Pediatr* 2021; 36: 566–573.

26. Wang Y and Wan CM. Guidelines for the diagnosis and treatment of acute fever without source in children 0 to 5 years of age in China (standard version). *Chin J Evid-Based Pediatr* 2008; 3: 449–457.

27. The taskforce of expert consensus on the implementation of the pediatric emergency triage system. Experts consensus on the implementation of the pediatric emergency triage system (Shanghai). *Chin Pediatr Emerg Med* 2022; 29: 849–857.

28. Lu XL. Treatment of high fever in children. *Chin J Appl Clin Pediatr* 2018; 33: 1388–1391.

29. Chen W, Rabhi F, Liao W, et al. Leveraging state-of-the-art topic modeling for news impact analysis on financial markets: a comparative study. *Electronics (Basel)* 2023; 12: 2605.

30. Harvey HB and Sotardi ST. The Pareto principle. *J Am Coll Radiol* 2018; 15: 931.

31. Van de Maat JS, Van Klink D, Den Hartogh-Griffioen A, et al. Development and evaluation of a hospital discharge information package to empower parents in caring for a child with a fever. *BMJ Open* 2018; 8: e021697.

32. Thompson AP, Le A, Hartling L, et al. Fading confidence: a qualitative exploration of parents' experiences caring for a febrile child. *J Clin Nurs* 2020; 29: 964–973.

33. Pan ZH and Ya MQIJH. Research on the current situation of children's medical care and the problems and countermeasures of carrying out hierarchical diagnosis and treatment in China. *Chin Gen Med* 2018; 21: 1177–1182.

34. Gu ST, Wang S, Wei MY, et al. Construction and application of children's internet hospital. *Acad J Second Military Med Univ* 2020; 41: 929–934.

35. Di T, Yang JY, Liu MD, et al. Research on development mode of internet hospital in China. *Chin Hosp* 2023; 27: 81–83.

36. Zhang R, Shao Y, Fan MQ, et al. Study on countermeasures for the development of internet medical services from the perspective of patient satisfaction. *J Med Inform* 2023; 44: 36–40.

37. Urbane UN, Likopa Z, Gardovska D, et al. Beliefs, practices and health care seeking behavior of parents regarding fever in children. *Medicina (B Aires)* 2019; 55: 398.

38. Yao M, Qi JG, Yan H, et al. Investigation and analysis of service status of general pediatric outpatient department of a third-grade class-a hospital in Beijing. *Chin Gen Med* 2015; 18: 1288–1292.

39. Wu XF. A study on the current status survey and influencing factors of parents' perception of fever and medical behavior of preschool children. Master Thesis, China: Qingdao University, 2021.

40. Yan Z, Hu WJ, Zhang JW, et al. Medical service status of pediatric outpatient and emergency department of a tertiary hospital in Beijing during COVID-19 epidemic: implications for the future. *Chin J Clinicians(Electronic Edition)* 2020; 14: 487–493.

41. Monsma J, Richerson J and Sloand E. Empowering parents for evidence-based fever management: an integrative review. *J Am Assoc Nurse Pract* 2015; 27: 222–229.