PERINATAL JOURNAL

Original Article

Perinatal Journal 2026; 34(2):86-100

https://doi.org/10.57239/prn.26.0342009

Leveraging sentiment analysis to assess hospital service quality and patient satisfaction: A systematic review

Felicia Hananto[1], Revi Margareta[2], Dewi Sri Surya Wuisan[3]*

[1,2,3]Magister of Hospital Management, Faculty of Medicine, Universitas Pelita Harapan, Indonesia

Abstract

Patient satisfaction is an important indicator in assessing the quality of hospital services and the success of the implementation of patient-centered care. Along with the development of digital technology, patients are increasingly expressing their experiences and perceptions through free texts on online platforms and social media. This condition opens up opportunities to use sentiment analysis as an alternative approach to evaluate the quality of health services more comprehensively and in real time. This study aims to systematically examine the use of the sentiment analysis method in assessing patient experience and satisfaction in health services. The method used is a systematic review with reference to the PRISMA 2020 guidelines. A literature search was conducted on the Scopus database for the period 2020–2025 and yielded 30 articles that met the inclusion criteria. The results of the study show that there is a significant methodological development from a lexicon-based approach towards machine learning, deep learning, and Large Language Models (LLMs). Thematically, patients' positive sentiments are mainly influenced by the quality of communication, empathy of healthcare workers, and interpersonal interactions, while negative sentiments are dominated by non-clinical factors such as wait times, administrative processes, and facilities. The conclusion of this study confirms that sentiment analysis is an effective and potential tool to support service quality evaluation, managerial decision-making, and strengthen patient-centered care approaches, although clinical validation and more contextual model development are still needed.

Keywords: Sentiment analysis, Patient satisfaction, Hospital service quality, Patient-Centered care

Introduction

Patient satisfaction has long been recognized as a critical indicator of hospital service quality and overall healthcare performance. As healthcare systems increasingly emphasize value-based care, patient satisfaction is no longer viewed merely as an outcome of service delivery, but as an integral component of quality assessment and continuous improvement. High levels of patient satisfaction are associated with better clinical outcomes, improved patient adherence to treatment, reduced complaints, and enhanced trust between patients and healthcare providers (AlOmari & Hamid, 2022). Moreover, patient satisfaction directly influences hospital reputation, public perception, and competitiveness, particularly in an era where patients have greater freedom to choose healthcare providers.

In the context of patient-centered care, understanding patients' experiences and perceptions is essential. Patient-centered care prioritizes individual needs, preferences, and values, ensuring that patients' voices are actively incorporated into healthcare decision-making processes (Balogun et al, 2024). Consequently, measuring patient satisfaction plays a strategic role in healthcare quality improvement initiatives, hospital accreditation, and policy formulation. Hospitals that effectively monitor and respond to patient feedback are more likely to enhance service quality, strengthen public trust, and maintain long-term sustainability.

Despite its importance, accurately assessing hospital service quality and patient satisfaction remains a significant challenge. Traditional approaches to measurement largely rely on structured surveys and questionnaires, such as patient satisfaction surveys or standardized instruments like the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) (Mercier et al, 2021). While these tools provide valuable quantitative insights, they have notable limitations. Survey-based methods are often conducted periodically rather than in real time, involve relatively small and selective respondent groups, and are susceptible to response bias, recall bias, and social desirability bias (Velykoivanenko et al, 2024). In addition, designing, administering, and analyzing surveys can be time-consuming and costly, limiting their practicality for continuous monitoring

Correspondence: Dewi Sri Surya Wuisan, Magister of Hospital Management, Faculty of Medicine, Universitas Pelita Harapan, Indonesia, **e-mail:** Dewi.wuisan@uph.edu, **Received:** November 21, 2025 **Accepted:** January 22, 2026

of service quality.

In recent years, patient behavior has shifted significantly in the digital era. Patients increasingly express their opinions, emotions, and experiences through online platforms such as Google Reviews, Twitter (X), Facebook, healthcare forums, and other social media channels. These online reviews and user-generated content represent rich sources of unsolicited, spontaneous, and natural patient feedback. Unlike structured surveys, online reviews capture patients' authentic perceptions of hospital services, including staff behavior, waiting times, facilities, communication quality, and overall care experiences. However, the sheer volume, unstructured nature, and linguistic variability of such textual data make manual analysis impractical and inefficient.

To address these challenges, sentiment analysis has emerged as a promising methodological approach. Sentiment analysis is a Natural Language Processing (NLP) technique used to automatically identify and classify opinions expressed in textual data as positive, negative, or neutral (Khaiser et al, 2023). By leveraging machine learning and linguistic algorithms, sentiment analysis enables rapid and scalable analysis of large volumes of patient-generated text data. This technique has been increasingly applied to assess patient satisfaction, evaluate service quality, analyze patient experiences, and examine public perceptions of healthcare institutions.

The application of sentiment analysis in hospital service evaluation offers significant advantages for healthcare management and quality improvement. It allows hospital administrators to monitor patient feedback in near real time, identify emerging issues, and detect patterns of dissatisfaction that may not be captured through traditional surveys. Furthermore, insights derived from sentiment analysis can support evidence-based decision-making, guide service redesign, enhance patient-centered care strategies, and ultimately improve the overall quality of hospital services.

Methods

This study employed a systematic review approach to examine the application of sentiment analysis methods in evaluating patient experiences and satisfaction within healthcare services. The review process was conducted in a structured and transparent manner in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines, ensuring methodological transparency, reproducibility, and scientific validity. All stages of the literature search, study selection, eligibility assessment, and final inclusion were summarized using a PRISMA flow diagram.

The literature search was conducted in August 2025 using the Watase UAKE scientific database, which is integrated with the Scopus database. The search was performed using the primary keyword "Patient Sentiments", with publication years restricted to 2020–2025. This time frame was selected to capture the most recent developments in patient sentiment analysis research, particularly in light of rapid advances in Natural Language Processing (NLP), machine learning, deep learning, and Large Language Models (LLMs). Only articles published in reputable peer-reviewed journals, spanning quartiles Q1 to Q4, were considered.

The initial search yielded 114 articles deemed potentially relevant based on keyword matching and available metadata.

During the identification stage, duplicate records, articles that did not meet the publication year criteria, and records with administrative or indexing issues were excluded. Following this step, 78 articles remained for the screening phase.

The screening stage involved a review of titles and abstracts to assess alignment with the study focus, namely the use of sentiment analysis to evaluate patient experiences within healthcare contexts. At this stage, 24 articles were excluded because they were not related to the healthcare domain, did not adopt a patient-centered perspective, or did not employ sentiment analysis techniques. Consequently, 54 articles were deemed eligible for full-text review.

The full-text review stage involved a comprehensive assessment of the remaining 54 articles to evaluate methodological rigor and the availability of systematically extractable information. Eligibility assessment focused on the clarity of sentiment analysis methodology, the presence of empirical data, and the relevance of findings to the study objectives. At this stage, 15 articles were excluded for the following reasons:

1. Insufficient description of sentiment analysis methods;
2. Absence of analyzable empirical results;
3. Lack of direct relevance to patient experience or satisfaction; and
4. Incomplete methodological information that hindered consistent data extraction.

Accordingly, 39 articles were deemed to meet the initial eligibility criteria.

To broaden the scope of the review and enrich the analytical perspective, additional relevant studies were identified from sources outside the Watase UAKE scientific database, applying the same publication period and journal quality criteria. Following the consolidation and re-verification process, a total of 30 articles were considered at the final inclusion stage.

These 30 articles were subsequently designated as the included studies in this systematic review. All included studies provided sufficient and consistent methodological information to allow systematic data extraction, encompassing:

1. Healthcare domain and service context;
2. Sources and characteristics of textual data;
3. Preprocessing procedures;
4. Feature extraction methods;
5. Sentiment analysis algorithms or models employed; and
6. Key research findings.

Data from the 30 included studies were analyzed qualitatively using a narrative synthesis approach. The analysis focused on identifying methodological trends, comparing sentiment analysis approaches (lexicon-based, machine learning, deep learning, and Large Language Model–based methods), and examining the implications of the findings for evaluating patient experience and satisfaction in healthcare services. This approach enabled a comprehensive understanding of the evolution and contribution of sentiment analysis in supporting healthcare quality improvement and patient-centered care.

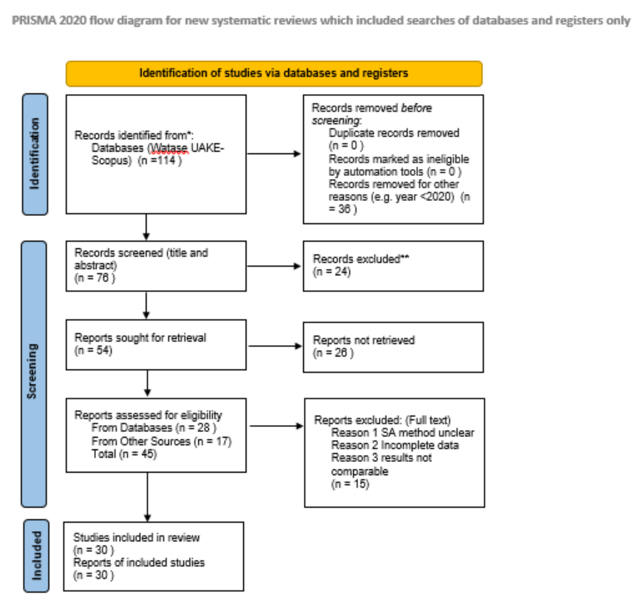

Figure 1. PRISMA flow

Result

Table 1: Included studies

No	Author	Year	Title	Research Objectives	Analysis Method	Tools/Model/Algorithm	Key Findings
1	El-Badaoui et al.	2025	A Novel Ontology-Based Approach for Analyzing Patient Sentiment Regarding	Ontology-based analysis of chronic disease patient sentiment	Ontology-based SA	Ontology engine, NLP	High accuracy

			Chronic Diseases				
2	Azarpey et al.	2025	NLP of Sentiments Identified in Patient Comments ... Less Than Top-Rated Care	Identify the negative sentiment of low-rated services	NLP SA	Rule-based + ML	Poor service factors detected
3	Murray et al.	2025	Probabilistic emotion and sentiment modelling ... patient-reported experiences	Probabilistic model of patient emotions	Probabilistic modelling	Bayesian model	The patient's emotions are clearly mapped
4	Sheng & Zheng	2025	Relationship between the sentiment of nursing notes ... mortality	The relationship between nurse record sentiment and mortality	Sentiment extraction	ML classifier	Negative sentiment predictors of mortality
5	Chatzimina et al.	2025	Topic modeling and sentiment analysis ... malignancies	Analysis of cancer patients' clinical conversations	LDA+SA	LDA, Lexicon	Empathy increases positive sentiment
6	Yuan et al.	2025	Identifying Patients' Preferences During Hospital Experience	Patient preferences from online comments	LDA+SA	ML+LDA	Hygiene & communication are dominant
7	Li et al.	2025	Revealing patient dissatisfaction ... ICD-11	Patient dissatisfaction from LLM+ICD11	Aspect SA + LLM	LLM, ICD-11	Dominant dissatisfaction with service availability
8	Nichini et al.	2025	Characterizing patient experience of physical restraint...	Patient experience with physical restraint	SA + Metaphor analysis	NLP metaphor engine	Restraint causes significant trauma
9	Kornblith et al.	2025	Analyzing patient perspectives with LLMs	LLM analysis of informed consent exceptions	Thematic + SA	LLM	Lack of information in patients
10	Dai et al.	2024	Medical service quality evaluation ... 7 chronic diseases	Evaluation of the quality of chronic disease services	LDA+SA	LDA, ML classifier	Lead times & costs are the most dominant
11	Al-Hadhrami et al.	2024	Deep learning sentiment analysis for drug reviews	Sentiment analysis of drug reviews	Deep learning SA	LSTM, CNN	Dominant side effects
12	Walsh et al.	2024	Combining topic modeling, SA,	Modafinil patient experience	Topic+SA	LDA, LIWC	High anxiety & side effects

			corpus linguistics...				
13	Chen et al.	2024	Analyzing patient experience on Weibo	Analysis of the patient experience via Weibo	Topic+SA	SVM, LDA	Lead time & communication are important
14	AlNasser & AlMuhaideb	2024	Advanced Arabic Aspect-Based Sentiment Analysis...	ABSA for Arab patient comments	ABSA	Transformers	High accuracy of models
15	Mohammad et al.	2024	Optimizing LLMs for Arabic Healthcare Communication	Optimization of LLMs for health communication	LLM optimization	Arabic LLM	LLM improves understanding of the Arabic context
16	Alkhnbashi et al.	2024	Aspect-based sentiment analysis of patient feedback using LLMs	ABSA-based LLM	ABSA	LLM	An accurate model in the service aspect
17	Gondode et al.	2024	Evaluation of AI-generated leaflets: ChatGPT vs Gemini	Evaluation of readability & sentiment on AI leaflets	Readability+Sentiment	GPT vs Gemini	Gemini is more accurate, ChatGPT is more humane
18	Sorayaie Azar et al.	2024	Predicting patients' sentiments about medications	Prediction of patient sentiment about drugs	AI prediction	ML+Deep learning	Accurate prediction of side effects
19	Srivastava et al.	2025	Graph neural network for sentiment polarity of COVID patients	GNN for COVID sentiment identification	GNN sentiment	Graph Neural Network	Effective GNN model
20	Pandey et al.	2023	Advanced sentiment analysis for patient experience...	SA for GP patient experience	Advanced SA	ML models	GP systems can be classified precisely
21	Yazdani et al.	2023	Sentiment analysis for Persian cancer patients	Sentiment of inpatient cancer patients	SA	Persian NLP	Dominant negative emotions during hospitalization
22	Khan et al.	2023	Hybrid market offering in the medical tech sector	Hybrid assessment offers medtech	Thematic + SA	NLP	Consumer sentiment towards mixed technologies
23	Tang et al.	2023	Sentiment Analysis of Patient Reviews of	Analyze factors that influence patient review	Lexicon-based sentiment analysis	VADER, unigram & bigram	Younger doctors get more positive reviews; Female

			Spine Surgeons	sentiment against spine surgeons			doctors are more likely to get lower ratings; Non-clinical factors are highly influential.
24	Guerrero et al.	2022	Detection of Hospital Patient Emotions from Online Reviews	Identifying and classifying patients' emotions from hospital reviews	Deep learning & machine learning	BiGRU–MCCNN, GloVe, FastText, PubMed+PMC	BiGRU–MCCNN model with PubMed+PMC embedding provides the best and most stable performance.
25	Khanbhai et al.	2022	Analysis of Patient Experience on Treatment Transition	Identify patient experience issues during the care transition	Machine learning	SVM, Bag of Words	NLP effectively detects problems such as a lack of information and poor discharge processes
26	Fan et al.	2022	The Influence of Online Comment Sentiment on Patient Choice	Analyze the impact of online comments on healthcare selection decisions	Lexicon-based sentiment analysis	SnowNLP	The quality of doctor-patient communication is the main factor in positive sentiment and patient choice.
27	Juan et al.	2022	Emotional Analysis of COVID-19 Patients from Electronic Nursing Records	Distinguishing the emotions of COVID-19 and non-COVID-19 patients	Lexicon-based & machine learning	NRC, AFINN, Bing, Neural Network	Electronic nursing records are effective as real-world data for patient emotional analysis.
28	Rahim et al.	2021	Evaluation of Hospital Service Quality Using Social Media	Evaluate the quality of public hospital services through Facebook reviews	Machine learning	SVM, TF-IDF	Social media is effective for measuring the quality of service and identifying areas of improvement.
29	Nawab et al.	2020	Identifying Negative Experiences of Inpatients	Identifying the main factors of the negative experience of hospital patients	Neural network	Neural Network, Bag of Words	NLP can uncover the main causes of negative experiences of inpatients

30	Khaleghparast et al.	2023	Classification of Cardiovascular Patient Satisfaction Using a Hybrid Approach	Classifying cardiovascular patient satisfaction	Hybrid lexicon & machine learning	TF-IDF, BoW, Multi-Feature Lexicon, MNB	A hybrid approach improves the accuracy of patient satisfaction classification.

Based on the 30 included studies, patient sentiment analysis in the healthcare domain demonstrates substantial advancement in terms of methodological approaches, data sources, and applied objectives. Temporally, the reviewed studies span the period 2020–2025, with a marked increase in publication intensity in recent years, coinciding with rapid progress in Natural Language Processing (NLP), deep learning, and Large Language Models (LLMs). These studies originate from diverse geographical and cultural contexts, including the Americas, Europe, Asia, and the Middle East, indicating that patient sentiment analysis has evolved into a globally adopted approach for evaluating healthcare experiences and service quality (Rahim et al., 2021; Tang et al., 2023; Guerrero et al., 2022; El-Badaoui et al., 2025).

All reviewed studies consistently utilized free-text data as the primary source of analysis. Such data were derived from a wide range of channels, including patient reviews on online platforms and hospital social media, patient experience surveys, comments on physician rating websites, as well as electronic health records and nursing notes. This diversity of data sources reflects a paradigm shift in healthcare service evaluation—from reliance on structured survey instruments toward the utilization of spontaneous, contextualized, and semantically rich patient narratives (Nawab et al., 2020; Juan et al., 2022; Khanbhai et al., 2022). These findings reinforce the view that free-text data can capture nuanced dimensions of patient experience that are often overlooked by conventional quantitative instruments.

From a data processing perspective, most studies reported the implementation of preprocessing procedures prior to sentiment analysis. These procedures typically included text cleaning to remove non-alphabetic characters, case normalization, punctuation removal, and language-specific linguistic adjustments. Although preprocessing steps were not always reported in a standardized manner, their application was widely regarded as essential for improving the consistency and quality of textual representations, particularly when dealing with unstructured data from social media and online comments (Rahim et al., 2021; Khaleghparast et al., 2023).

At the stage of feature extraction and text representation, a variety of approaches were observed, reflecting the methodological evolution of the field. Early and medium-scale studies predominantly employed traditional techniques such as Bag of Words and TF-IDF, which proved effective for datasets with limited complexity (Rahim et al., 2021; Khanbhai et al., 2022). Subsequent research increasingly adopted word embedding techniques—including GloVe, FastText, and PubMed+PMC—to capture richer semantic relationships within medical texts (Guerrero et al., 2022). More recent studies demonstrate a clear shift toward contextual representations based on Transformer architectures and LLMs, particularly in research focusing on Aspect-Based Sentiment Analysis (ABSA), multilingual analysis, and integration with clinical taxonomies such as ICD-11 (Li et al., 2025; AlNasser & AlMuhaideb, 2024).

From the perspective of analytical methods, the approaches employed across the 30 studies can be categorized into several main groups: lexicon-based analysis, classical machine learning, hybrid approaches, deep learning, topic modeling, and LLM-based methods. Lexicon-based techniques—such as VADER, NRC, AFINN, and Bing—were primarily applied in exploratory studies or those involving relatively small datasets (Tang et al., 2023; Juan et al., 2022). In contrast, classical machine learning algorithms, including Support Vector Machine (SVM), Naïve Bayes, and Logistic Regression, were widely used for sentiment classification and demonstrated satisfactory performance within healthcare contexts (Rahim et al., 2021; Khanbhai et al., 2022).

Hybrid approaches combining lexicon-based features with machine learning models consistently outperformed single-method approaches, particularly in addressing linguistic variability and class imbalance (Khaleghparast et al., 2023). Furthermore, deep learning architectures—such as CNN, LSTM, GRU, and BiGRU–MCCNN—exhibited superior performance in detecting complex emotions and sentiment patterns, especially when paired with domain-specific medical embeddings (Guerrero et al., 2022). These findings underscore the importance of aligning model complexity with data characteristics and analytical objectives.

Thematically, all studies demonstrated that sentiment analysis is effective in identifying patient perceptions, satisfaction levels, and emotional responses to healthcare services. Cross-study synthesis revealed relatively consistent patterns regarding factors influencing patient sentiment. Positive sentiment was commonly associated with effective communication between healthcare professionals and patients, empathetic attitudes, clarity of information, and interpersonal interactions that respect patients as active participants in care (Rahim et al., 2021; Chatzimina et al., 2025; Tang et al., 2023). Conversely, negative sentiment was most frequently triggered by long waiting times, limited service availability, administrative barriers, and insufficient information during care processes (Nawab et al., 2020; Khanbhai et al., 2022; Dai et al., 2024).

Notably, multiple studies emphasized that non-clinical experiences make a substantial contribution to patient sentiment, often exerting a stronger influence than clinical aspects alone. Factors such as facility cleanliness, hospital environment, noise levels, ease of registration and discharge procedures, and service costs consistently emerged as key determinants of patient sentiment (Tang et al., 2023; Chen et al., 2024; Dai et al., 2024). These findings suggest that patients perceive healthcare quality holistically, extending beyond clinical outcomes.

Another significant development identified in this synthesis is the expansion of sentiment analysis from a purely evaluative tool to a predictive and clinically relevant instrument. Sheng and Zheng (2025) demonstrated that negative sentiment expressed in nursing notes is associated with increased mortality risk, highlighting the potential of sentiment analysis as an early warning system in clinical care. This finding aligns with the probabilistic modeling approach proposed by Murray et al. (2025), which revealed that clinical narratives contain latent, clinically meaningful emotional information.

In addition, studies focusing on chronic diseases, pharmacovigilance, and drug reviews showed that sentiment analysis can detect complaint patterns and adverse drug effects more rapidly than traditional reporting systems (El-Badaoui et al., 2025; Al-Hadhrami et al., 2024; Sorayaie Azar et al., 2024). These results reinforce the potential of sentiment analysis as a supportive tool for clinical decision-making and drug regulation, although challenges related to subjectivity and data irregularity remain.

Approaches integrating topic modeling and conversational analysis further demonstrated that sentiment cannot be interpreted independently of topical context. Studies by Chatzimina et al. (2025) and Walsh et al. (2024) indicated that negative sentiment does not necessarily reflect dissatisfaction with services, but may instead represent severe clinical conditions or emotional burden. Consequently, integrating topic modeling with sentiment analysis is crucial for maintaining contextual and ethical interpretation.

At the same time, the use of LLMs and Transformer-based models presents both opportunities and challenges. Studies by Li et al. (2025) and Alkhnbashi et al. (2024) showed that LLMs improve the accuracy of aspect-level analysis and contextual understanding, including in Arabic-language and specialized medical domains. However, Kornblith et al. (2025) cautioned against risks of algorithmic bias and hallucination, particularly in sensitive contexts such as medical ethics and informed consent. This underscores the necessity of expert validation in the application of LLMs within healthcare.

Overall, the synthesis of these 30 studies demonstrates that free-text–based sentiment analysis has evolved from an exploratory technique into a strategic tool for healthcare service evaluation, clinical monitoring, pharmacovigilance, and health policy formulation. Nevertheless, key challenges persist, including linguistic and cultural bias, dataset imbalance, limited cross-context generalizability, and the need to integrate advanced models with clinical verification. Addressing these challenges will enable sentiment analysis to more effectively strengthen patient-centered care and support evidence-based decision-making in the healthcare sector.

Discussion

The findings of this systematic review demonstrate a clear and progressive methodological evolution in patient sentiment analysis within the healthcare domain. This evolution spans from relatively simple lexicon-based approaches to machine learning, hybrid methods, deep learning, and, more recently, the adoption of Large Language Models (LLMs). This shift reflects a growing need to capture patient narratives that are increasingly complex, contextual, and multidimensional, particularly as digital platforms have become a dominant medium through which patients express their healthcare experiences.

Studies employing lexicon-based approaches tend to emphasize their advantages in terms of efficiency and ease of implementation. These methods do not require data training or annotation, making them relatively straightforward to apply to small- to medium-sized datasets (Tang et al., 2023; Fan et al., 2022; Juan et al., 2022). Consequently, lexicon-based techniques have been widely used in early exploratory research to obtain an initial overview of patient sentiment toward healthcare services. However, their limited ability to capture linguistic context, irony, metaphor, and the emotional complexity of patient expressions represents a major drawback, particularly in the healthcare domain, which is characterized by technical terminology and emotionally ambiguous language.

The application of machine learning methods has demonstrated a notable improvement in performance compared with lexicon-based approaches, especially when applied to larger and more heterogeneous datasets. Several studies reported that Support Vector Machine (SVM) is among the most stable and consistent algorithms for sentiment classification in medical text (Rahim et al., 2021; Khanbhai et al., 2022). This observation is further supported by Zhang et al. (2025), who showed that SVM exhibits strong discriminative power and generalization ability across repeated evaluations on healthcare text data. The strength of SVM lies in its ability to handle high-dimensional feature spaces and construct optimal decision boundaries, making it a reliable choice for patient sentiment analysis in small- to medium-scale studies.

Subsequently, hybrid approaches combining lexicon-based features with machine learning algorithms have shown even greater performance improvements. Khaleghparast et al. (2023) demonstrated that integrating lexicon-derived features with machine learning models resulted in substantial gains in classification accuracy and F1-score. This finding aligns with Sham and Mohamed (2022), who argued that hybrid approaches effectively leverage the complementary strengths of both methods—namely, the interpretability of lexicons and the predictive power of machine learning. These results indicate a methodological trend toward more integrative and adaptive approaches in interpreting patient sentiment.

The most prominent methodological advancement is reflected in the increasing use of deep learning techniques, particularly in studies involving large and complex datasets. Guerrero et al. (2022) showed that the BiGRU–MCCNN model outperformed traditional approaches not only in terms of accuracy but also in its stability when detecting infrequent emotional expressions. This capability is highly relevant in healthcare contexts, where subtle emotions such as anxiety, fear, or frustration may carry significant clinical implications. These findings are consistent with Guo (2021), who reported exceptionally high accuracy in human emotion detection using deep learning models applied to large-scale textual data. Collectively, these results confirm that deep learning models are better equipped to handle the complexity of natural language and the diversity of patient emotions compared with conventional methods.

In more recent studies, the use of Transformer architectures and LLMs has become increasingly prominent, particularly in the context of Aspect-Based Sentiment Analysis (ABSA), multilingual analysis, and integration with clinical terminology. Li et al. (2025) demonstrated that LLMs combined with ICD-11–based classification can identify patient dissatisfaction aspects in a more detailed and systematic manner. Similar conclusions were reported by AlNasser and AlMuhaideb (2024), Mohammad et al. (2024), and Alkhnbashi et al. (2024), all of whom highlighted the superior capability of Transformer-based and LLM approaches in understanding linguistic context, especially in Arabic-language and complex medical domains. Nevertheless, Kornblith et al. (2025) cautioned against the risks of algorithmic bias and hallucination associated with LLMs, particularly in sensitive contexts such as medical ethics and informed consent. This underscores the critical importance of expert validation in the application of LLMs in healthcare settings.

From a thematic perspective, the synthesis reveals strong consistency regarding the factors influencing patient sentiment. Positive sentiment is predominantly driven by the quality of interpersonal interactions between healthcare professionals and patients, including clear communication, empathy, friendliness, and respect for patients as individuals (Tang et al., 2023; Fan et al., 2022; Chatzimina et al., 2025). Effective interpersonal communication fosters a sense of safety, trust, and understanding, which patients subsequently express through positive sentiment. Non-verbal elements such as tone of voice, eye contact, and body language further shape patients' emotional perceptions and contribute to their overall care experience (Heyn et al., 2022).

In contrast, negative sentiment is more frequently associated with non-clinical factors, such as facility cleanliness, noise levels, long waiting times, service costs, and administrative and discharge procedures (Nawab et al., 2020; Dai et al., 2024; Chen et al., 2024). These findings indicate that patient perceptions of healthcare quality are holistic and not solely determined by clinical outcomes. Neglecting non-clinical aspects may significantly undermine patient satisfaction, even when the quality of medical treatment itself is high.

The practical implications of this systematic review are particularly relevant for hospital management and policymakers. Sentiment analysis enables near real-time monitoring of patient satisfaction by leveraging reviews distributed across multiple digital platforms. By systematically analyzing this data, healthcare institutions can detect declines in service quality at an early stage, design more proactive interventions, and mitigate potential reputational damage (Pandey et al., 2023; Yuan et al., 2025). Moreover, several studies indicate that sentiment analysis can be extended beyond evaluation toward predictive applications, such as monitoring clinical risk and patient mortality (Sheng & Zheng, 2025; Murray et al., 2025), thereby opening opportunities for integration into clinical decision support systems.

Despite these advantages, this review also identifies several persistent methodological challenges. Most studies rely on single data sources and relatively small datasets, increasing susceptibility to selection bias and limited generalizability (Guerrero et al., 2022; Khanbhai et al., 2022; Tang et al., 2023; Rahim et al., 2021). Additionally, sentiment class imbalance is a commonly reported issue that can adversely affect model performance (Khaleghparast et al., 2023). Juan et al. (2022) further highlighted the lack of suitable benchmark datasets due to the relative novelty of this research area.

Accordingly, future research is encouraged to utilize multi-source and multilingual datasets, apply techniques for handling imbalanced data such as down-sampling, up-sampling, and synthetic data generation, and develop sentiment analysis models that are more adaptive and clinically validated. Integrating advanced approaches such as deep learning and LLMs with clinical domain expertise is essential to ensure that sentiment analysis is not only technically accurate but also meaningful, ethical, and actionable within healthcare practice.

Limitations and research gaps

Although research on patient sentiment analysis in healthcare has expanded rapidly over the past five years, existing studies continue to exhibit several methodological and conceptual limitations that give rise to important research gaps requiring further investigation.

First, the majority of studies still rely heavily on online review data, such as Google Reviews, Twitter, or healthcare forums (El-Badaoui, 2021; Azarpey, 2020; Li, 2021; Yuan, 2022). These data sources are inherently voluntary, non-standardized, and biased toward extreme experiences—either highly positive or highly negative. Consequently, such reliance introduces selection bias and representativeness bias, which limit the generalizability of findings to the broader patient population. Only a limited number of studies have integrated public online reviews with Patient-Reported Experience Measures (PREMs) or structured clinical records, resulting in an incomplete and non-holistic representation of patient sentiment.

Second, most sentiment analysis models currently employed—including Aspect-Based Sentiment Analysis (ABSA) and topic modeling—remain insufficient for capturing deep clinical context, particularly in relation to irony, ambiguity, and emotionally nuanced patient language. Although recent Transformer-based models and LLMs have demonstrated improvements in contextual understanding (Al-Hadhrami, 2023; AlNasser, 2022), numerous studies caution that algorithmic bias and limited domain adaptation may lead to misinterpretation of sensitive medical language. This gap highlights the need to develop clinically validated LLMs, specifically tailored to the healthcare domain and systematically tested across cultural and linguistic contexts.

Third, several studies that utilize sentiment data for clinical prediction, such as mortality or prognosis assessment (Sheng & Zheng, 2021), demonstrate considerable potential but remain constrained by data drawn predominantly from single institutions or isolated electronic health record systems. This limitation raises concerns regarding external validity. A significant research gap persists in determining how sentiment analysis can be safely, standardly, and interoperably integrated into early warning systems or clinical decision support systems (CDSSs).

Fourth, while some studies have addressed pharmacovigilance and the detection of adverse drug reactions (Sorayaie Azar, 2022; Al-Hadhrami, 2023), most research continues to focus on general patient perceptions rather than applying sentiment analysis to specific clinical evaluations, such as chronic disease management, inpatient care experiences, or emergency services. This indicates a substantial research gap in high-risk and clinically sensitive settings, where patient sentiment may carry critical implications.

Fifth, there remains a scarcity of research examining the influence of social, cultural, and psychological factors on sentiment expression, particularly in contexts such as mental health, traumatic experiences, or invasive medical procedures (Nichini, 2021; Kornblith, 2022). Most existing models implicitly assume that sentiment expression is universal, despite strong evidence that emotional expression is deeply shaped by cultural background and individual experiences. This gap underscores the need for interdisciplinary approaches that integrate healthcare research with linguistics and psychology.

Finally, there is a notable lack of studies investigating how sentiment analysis results are operationally utilized by hospitals or regulators to drive tangible service improvements. Most existing research remains descriptive or exploratory, rather than evaluative or interventional. This gap raises a critical question: how can sentiment analysis be effectively translated into operational policies, hospital standard operating procedures (SOPs), clinical communication interventions, or tangible improvements in queue management and service delivery systems?

Conclusion

This review demonstrates that patient sentiment analysis has evolved into an increasingly strategic approach for evaluating service quality and patient experience in healthcare. Based on the synthesis of 30 included studies and supporting literature, it can be concluded that the use of free-text data derived from online reviews, social media, surveys, and clinical records is capable of revealing patient perceptions, satisfaction, and emotional responses that are not consistently captured by structured survey instruments. Cross-study findings consistently indicate that non-clinical factors—such as waiting time, communication with healthcare professionals, facility cleanliness, and the ease of administrative processes—exert a significant influence on patient sentiment, often equaling or exceeding the impact of certain clinical aspects (Rahim et al., 2021; Tang et al., 2022; Nawab et al., 2020).

From a methodological perspective, the reviewed studies reveal a clear shift from lexicon-based and traditional machine learning approaches toward the adoption of deep learning and Large Language Models (LLMs) to capture more complex linguistic and emotional contexts. These advanced approaches have been shown to improve the accuracy of sentiment analysis, particularly in medical language and long-form narrative patient reviews (Al-Hadhrami et al., 2024; AlNasser & AlMuhaideb, 2024; Kornblith et al., 2025). Furthermore, several studies indicate that patient sentiment functions not only as an indicator of service perception but also holds clinical potential, such as predicting mortality risk and monitoring adverse drug reactions (Sheng & Zheng, 2025; Sorayaie Azar et al., 2024).

Overall, this review confirms that sentiment analysis represents a promising data-driven tool for supporting patient-centered care, healthcare quality improvement, and evidence-based decision-making. However, the extent to which its benefits can be fully realized remains highly dependent on data quality, the alignment of analytical models with linguistic and cultural contexts, and the effective integration of analytical outputs into real-world clinical and managerial practices.

Based on the findings, limitations, and research gaps identified in this review, several directions for future research are recommended. First, future studies should integrate sentiment data from online reviews with Patient-Reported Experience Measures (PREMs), Patient-Reported Outcome Measures (PROMs), and structured clinical data to generate a more comprehensive and representative understanding of patient experience. Such a multimodal approach has the potential to reduce bias and enhance external validity.

Second, the development and evaluation of NLP models—particularly LLMs—should be increasingly focused on the healthcare domain through clinical expert validation, cross-cultural testing, and local language adaptation. Models should not only achieve high technical accuracy but also meet standards of clinical safety, ethical accountability, and interpretability.

Third, future research should expand the application of sentiment analysis toward predictive and preventive functions, including early warning systems, chronic disease monitoring, and clinical decision support. Achieving this objective will require longitudinal study designs, multi-institutional datasets, and clearly defined interoperability frameworks.

Fourth, more in-depth research is needed in high-risk and sensitive healthcare contexts, such as mental health services, intensive care, invasive procedures, and patient trauma experiences. Integrating sentiment analysis with qualitative and psychosocial approaches would be particularly valuable in uncovering the deeper meanings underlying patients' emotional expressions.

Finally, future research should move beyond descriptive analyses toward implementative and interventional studies that evaluate how sentiment analysis findings are translated into service improvements, policy formulation, and changes in clinical practice. In doing so, sentiment analysis can evolve from a purely analytical technique into a practical instrument for enhancing the quality, resilience, and sustainability of healthcare systems grounded in patient-centered care.

References

A. Rahim, A. I., Ibrahim, M. I., Musa, K. I., Chua, S.-L., & Yaacob, N. M. (2021). Assessing Patient-Perceived Hospital Service Quality and Sentiment in Malaysian Public Hospitals Using Machine Learning and Facebook Reviews. International Journal of Environmental Research and Public Health, 18(18), 9912. https://doi.org/10.3390/ijerph18189912

Al-Hadhrami, S., Vinko, T., Al-Hadhrami, T., Saeed, F., & Qasem, S. N. (2024). Deep learning-based method for sentiment analysis for patients' drug reviews. PeerJ Computer Science, 10, e1976.

Alkhnbashi, O. S., Mohammad, R., & Hammoudeh, M. (2024). Aspect-based sentiment analysis of patient feedback using large language models. Big Data and Cognitive Computing, 8(12), 167.

AlNasser, S., & AlMuhaideb, S. (2024). Listening to

Patients: Advanced Arabic Aspect-Based Sentiment Analysis Using Transformer Models Towards Better Healthcare. Big Data and Cognitive Computing, 8(11), 156.

AlOmari, F., & A. Hamid, A. B. (2022). Strategies to improve patient loyalty and medication adherence in Syrian healthcare settings: The mediating role of patient satisfaction. PloS one, 17(11), e0272057.

Azarpey, A., Thomas, J., Ring, D., & Franko, O. (2025). Natural Language Processing of Sentiments Identified in Patient Comments Associated with Less Than Top-Rated Care. Journal of Patient Experience, 12, 23743735251323677.

Balogun, O. D., Mustapha, A. Y., Tomoh, B. O., Soyege, O. S., Nwokedi, C. N., Mbata, A. O., & Iguma, D. R. (2024). Patient-centered care models: A review of their influence on healthcare management practices. Journal of Frontiers in Multidisciplinary Research, 5(2), 28-35.

Chatzimina, M. E., Papadaki, H. A., Pontikoglou, C., & Tsiknakis, M. (2025). Topic modeling and sentiment analysis of Greek clinician–patient conversations in hematologic malignancies. International Journal of Medical Informatics, 106071.

Chen, X., Shen, Z., Guan, T., Tao, Y., Kang, Y., & Zhang, Y. (2024). Analyzing patient experience on Weibo: machine learning approach to topic modeling and sentiment analysis. JMIR Medical Informatics, 12, e59249.

Dai, J., Lyu, F., Yu, L., Zhou, Z., & He, Y. (2024). Medical service quality evaluation based on LDA and sentiment analysis: Examples of seven chronic diseases. Digital health, 10, 20552076241233864.

El-Badaoui, M., Gherabi, N., Qanouni, F., & Amnai, M. (2025). A Novel Ontology-Based Approach for Analyzing Patient Sentiment Regarding Chronic Diseases. International Journal of Online & Biomedical Engineering, 21(8).

Fan, J., Geng, H., Liu, X., & Wang, J. (2022). The Effects of Online Text Comments on Patients' Choices: The Mediating Roles of Comment Sentiment and Comment Content. Frontiers in Psychology, 13(886077). https://doi.org/10.3389/fpsyg.2022.886077

Gondode, P. G., Khanna, P., Sharma, P., Duggal, S., & Garg, N. (2024). End-of-life care patient information leaflets—A comparative evaluation of artificial intelligence-generated content for readability, sentiment, accuracy, completeness, and suitability: ChatGPT vs Google Gemini. Indian Journal of Critical Care Medicine: Peer-reviewed, Official Publication of Indian Society of Critical Care Medicine, 28(6), 561.

Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. Journal of Intelligent Systems, 31(1), 113–126. https://doi.org/10.1515/jisys-2022-0001

Heyn, L. G., Løkkeberg, S. T., Ellington, L., van Dulmen, S., & Eide, H. (2023). Understanding the role of positive emotions in healthcare communication – A realist review. Nursing Open, 10(6). https://doi.org/10.1002/nop2.1604

Juan Nicolás CUENCA-ZALDÍVAR, TORRENTE-REGIDOR, M., MARTÍN-LOSADA, L., César FERNÁNDEZ-DE-LAS-PEÑAS, Lidiane Lima FLORENCIO, Pedro Alexandre SOUSA, & PALACIOS-CEÑA, D. (2022). Exploring Sentiment and Care Management of Hospitalized Patients During the First Wave of the COVID-19 Pandemic Using Electronic Nursing Health Records: Descriptive Study. JMIR Medical Informatics, 10(5), e38308–e38308. https://doi.org/10.2196/38308

Khaiser, F. K., Saad, A., & Mason, C. (2023). Sentiment analysis of students' feedback on institutional facilities using text-based classification and natural language processing (NLP). Journal of Language and Communication, 10(1), 101-111.

Khaleghparast, S., Maleki, M., Hajianfar, G., Soumari, E., Oveisi, M., Golandouz, H. M., Noohi, F., Dehaki, M. G., Golpira, R., Mazloomzadeh, S., Arabian, M., & Kalayinia, S. (2023). Development of a patient satisfaction analysis system using machine learning and lexicon-based methods. BMC Health Services Research, 23(1), 1–12. https://doi.org/10.1186/s12913-023-09260-7

Khan, H., Arslan, A., Haapanen, L., Rodgers, P., & Tarba, S. Y. (2023). Hybrid market offering in the medical technology sector and the role of network configuration: an exploratory

assessment in both developed and emerging markets. International Marketing Review, 40(4), 612-635.
Khanbhai, M., Warren, L., Symons, J., Flott, K., Harrison-White, S., Manton, D., Darzi, A., & Mayer, E. (2022). Using natural language processing to understand, facilitate, and maintain continuity in patient experience across transitions of care. International Journal of Medical Informatics, 157, 104642. https://doi.org/10.1016/j.ijmedinf.2021.104642
Kornblith, A. E., Singh, C., Innes, J. C., Chang, T. P., Adelgais, K. M., Holsti, M., ... & Counts, C. R. (2025). Analyzing patient perspectives with large language models: a cross-sectional study of sentiment and thematic classification on exception from informed consent. Scientific reports, 15(1), 6179.
Li, J., Yang, Y., Mao, C., Pang, P. C. I., Zhu, Q., Xu, D., & Wang, Y. (2025). Revealing patient dissatisfaction with health care resource allocation in multiple dimensions using large language models and the International Classification of diseases 11th revision: aspect-based sentiment analysis. Journal of Medical Internet Research, 27, e66344.
Mercier, M. R., Galivanche, A. R., David, W. B., Malpani, R., Pathak, N., Hilibrand, A. S., ... & Grauer, J. N. (2021). Hospital Consumer Assessment of Healthcare Providers and Systems survey response rates are significantly affected by patient characteristics and postoperative outcomes for patients undergoing primary total knee arthroplasty. PLoS One, 16(9), e0257555.
Mohamad Sham, N., & Mohamed, A. (2022). Climate Change Sentiment Analysis Using Lexicon, Machine Learning, and Hybrid Approaches. Sustainability, 14(8), 4723. https://doi.org/10.3390/su14084723
Mohammad, R., Alkhnbashi, O. S., & Hammoudeh, M. (2024). Optimizing Large Language Models for Arabic Healthcare Communication: A Focus on Patient-Centered NLP Applications. Big Data and Cognitive Computing, 8(11), 157.
Murray, C., Mitchell, L., Tuke, J., & Mackay, M. (2025). Probabilistic emotion and sentiment modelling of patient-reported experiences. Artificial Intelligence in Medicine, 103178.
Nawab, K., Ramsey, G., & Schreiber, R. (2020). Natural Language Processing to Extract Meaningful Information from Patient Experience Feedback. Applied Clinical Informatics, 11(02), 242–252. https://doi.org/10.1055/s-0040-1708049
Nichini, C., Barattieri di San Pietro, C., Scalingi, B., Alecci, E., Toschi, L., Cavallotti, S., ... & Bambini, V. (2025). Characterizing the patient experience of physical restraint in psychiatric settings via a linguistic, sentiment, and metaphor analysis. Scientific Reports, 15(1), 2111.
Pandey, A. R., Seify, M., Okonta, U., & Hosseinian-Far, A. (2023). Advanced sentiment analysis for managing and improving patient experience: Application for general practitioner (GP) classification in Northamptonshire. International Journal of Environmental Research and Public Health, 20(12), 6119.
Serrano-Guerrero, J., Bani-Doumi, M., Romero, F. P., & Olivas, J. A. (2022). Understanding what patients think about hospitals: A deep learning approach for detecting emotions in patient opinions. Artificial Intelligence in Medicine, 128, 102298. https://doi.org/10.1016/j.artmed.2022.102298
Sheng, L., & Zheng, Q. (2025). Relationship between the sentiment of nursing notes and one-year mortality of elderly sepsis patients. PLoS One, 20(5), e0323686.
Sorayaie Azar, A., Babaei Rikan, S., Naemi, A., Bagherzadeh Mohasefi, J., & Wiil, U. K. (2024). Predicting patients' sentiments about medications using artificial intelligence techniques. Scientific Reports, 14(1), 31928.
Srivastava, S., Chakraborty, C., & Sarkar, M. K. (2025). A graph neural network-based machine learning model for sentiment polarity and behavior identification of covid patients. International Journal of Data Science and Analytics, 19(4), 687-696.
Tang, J. E., Arvind, V., Dominy, C., White, C. A., Cho, S. K., & Kim, J. S. (2022). How Are Patients Reviewing Spine Surgeons Online? A Sentiment Analysis of Physician Review Website Written Comments. Global Spine Journal, 13(8), 219256822110699.

https://doi.org/10.1177/21925682211069933

Velykoivanenko, L., Salehzadeh Niksirat, K., Teofanovic, S., Chapuis, B., Mazurek, M. L., & Huguenin, K. (2024, May). Designing a data-driven survey system: leveraging participants' online data to personalize surveys. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (pp. 1-22).

Walsh, J., Cave, J., & Griffiths, F. (2024). Combining topic modeling, sentiment analysis, and corpus linguistics to analyze unstructured web-based patient experience data: case study of modafinil experiences. Journal of Medical Internet Research, 26, e54321.

Yazdani, A., Shamloo, M., Khaki, M., & Nahvijou, A. (2023). Use of sentiment analysis for capturing hospitalized cancer patients' experience from free-text comments in the Persian language. BMC Medical Informatics and Decision Making, 23(1), 275.

Yuan, J., Chen, X., Yang, C., Chen, J., Han, P., Zhang, Y., & Zhang, Y. (2025). Identifying Patients' Preferences During Their Hospital Experience. A Sentiment and Topic Analysis of Patient-Experience Comments via Natural Language Techniques. Patient preference and adherence, 2027-2037.

Zhang, Y., Le, Lopez, N., & Phan, K. (2025). Comparative Study of Machine Learning Models for Textual Medical Note Classification. Computers, 15(1), 7–7. https://doi.org/10.3390/computers15010007