International Journal of Healthcare Sciences ISSN 2348-5728 (Online)
Vol. 13, Issue 1, pp: (610-631), Month: April 2025 - September 2025, Available at: www.researchpublish.com

Sentiment Analysis of Patient Feedback for Strategic Business Decisions in Hospitals

Dr. David Augustine Bull

DBA, Ph.D., MBA, M.Sc., BSc., PMP., CMHC.

American InterContinental University, Online

DOI: https://doi.org/10.5281/zenodo.16964956

Published Date: 27-August-2025

Abstract: **Background. Patient narratives are increasingly recognized as a valuable source of intelligence for healthcare quality improvement, yet their systematic integration into performance management remains limited. Sentiment analysis, particularly aspect-based models, offers opportunities to align patient feedback with strategic metrics. This study investigated the role of sentiment analysis in patient experience, operational performance, leadership decision-making, and organizational sustainability.**

Methods. A multi-method analytic design was applied to a sample of 50 hospitals, yielding 600 hospital-month observations across one calendar year. Patient feedback was collected from internal hospital surveys, complaint logs, and narrative comments submitted through electronic portals. Data were de-identified in accordance with HIPAA standards, with duplicate entries removed and missing cases handled via listwise deletion (<1% loss). Aspect-based sentiment analysis was conducted using a validated LLM-assisted model, producing domain-level polarity scores (communication, access, billing, discharge). These sentiment scores were linked to Centers for Medicare & Medicaid Services (CMS)-reported HCAHPS domain scores and hospital operational metrics, including no-show rates, call-center wait times, and length of stay. For leadership decision-making (RQ3), a survey of 120 hospital leaders assessed adoption, perceived usefulness, and decision-making effectiveness using validated scales consistent with the Technology Acceptance Model. RQ4 compared hospitals with governed, validated sentiment pipelines (n = 25) and those without governance (n = 25) using quarterly aggregated KPI improvements.

Results. Aspect-based sentiment scores were significantly associated with HCAHPS domain scores (average $r = .42$, $p < .001$), supporting measurement validity. Regression models demonstrated that sentiment scores predicted operational outcomes, including reduced no-show rates ($\beta = -3.21$, $p < .001$) and shorter call-center wait times ($\beta = -2.3$ minutes, $p < .001$). Mediation analysis confirmed that perceived usefulness mediated the effect of adoption on decision-making effectiveness (indirect effect = 0.21, 95% CI [0.09, 0.37]). Finally, hospitals with governed pipelines achieved significantly greater and more sustainable KPI improvements, including a −4.7% reduction in no-shows versus −1.9% in non-governed hospitals, $t(48) = 7.28$, $p < .001$.

Conclusions. Sentiment analysis operates at multiple levels of healthcare performance: validating patient experiences, predicting operational efficiency, enhancing leadership decision-making through perceived usefulness, and sustaining improvements through governance. The findings advance theory by extending the Technology Acceptance Model and advance practice by offering a roadmap for embedding sentiment analysis as a strategic tool for patient-centered, efficient, and sustainable healthcare delivery.

Keywords: **sentiment analysis; patient experience; operational performance; Technology Acceptance Model; healthcare leadership; governance.**

I. INTRODUCTION

Hospitals and health systems are increasingly overwhelmed with patient feedback from multiple channels, including HCAHPS surveys, patient portals, emails, social media, and public review sites. Evidence indicates that patient portals, while designed to improve access and engagement, have significantly increased clinician workload: nearly 68% of physicians report feeling overwhelmed by the sheer volume of portal messages (OncLive, 2023), and post-pandemic trends show that message volumes continue to rise, driving greater time spent in electronic health records (EHRs) and contributing to burnout (American Medical Association [AMA], 2023). This "patient portal messaging crisis" has been recognized as a strain on clinician workflows and a driver of after-hours work, further exacerbating stress and inefficiency (Zivin & Chwistek, 2025). At the same time, patient feedback is no longer confined to traditional surveys; it increasingly arrives via emails, text-based surveys, QR codes, review platforms, and social media, amplifying both the scale and complexity of the feedback that health systems must manage (Press Ganey, 2024). Converting this unstructured, multi-channel input into reliable and actionable signals for decision-making has therefore become a strategic imperative. In parallel, payers and regulators have expanded the visibility and salience of patient-experience metrics through public reporting and incentive programs, raising the business stakes of effectively understanding what patients actually say and feel.

Within the United States, the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey remains the nationally standardized benchmark for patient experience, with results publicly reported and tied to several Centers for Medicare & Medicaid Services (CMS) programs that influence payment and reputation, such as Star Ratings and Care Compare (Centers for Medicare & Medicaid Services [CMS], 2023). Consequently, health systems are increasingly working to align internal, real-time feedback sources, including comments, online reviews, and portal messages—with HCAHPS domains to guide operational priorities and service investments that directly impact both market share and reimbursement. A complementary stream of research has examined online patient reviews (OPRs) on platforms such as Google and Yelp. Findings suggest that review text and star ratings are correlated with dimensions of patient experience, though the evidence linking OPRs with clinical quality measures is mixed, implying that OPRs more reliably capture service perceptions than clinical outcomes (Hao et al., 2023; Ranard et al., 2016; Segal et al., 2018). For healthcare leaders, this indicates that OPR mining can be valuable for service recovery and brand management, while clinical quality must still be assessed through registries and internal performance data (Wang et al., 2023). Methodologically, advances in natural language processing (NLP) have enabled sentiment analysis to move beyond lexicon-based rules toward transformer-based models (e.g., BERT variants) and aspect-based sentiment analysis (ABSA), which assign polarity to specific experience themes such as communication, wait times, discharge processes, and billing. Recent healthcare applications demonstrate that ABSA and transformer models can reveal granular drivers of satisfaction and dissatisfaction, enabling targeted interventions such as triage staffing adjustments, discharge education redesign, or billing process improvements (Li et al., 2022; Yadav et al., 2021). Comparative studies also show that while traditional lexicon-based tools (e.g., VADER) remain accessible, transformer models such as BERT and Flair generally outperform them when sufficient labeled data are available (Karisani & Karisani, 2021; Zhou et al., 2022).

Strategically, healthcare organizations are beginning to link sentiment data streams to operational performance indicators such as emergency department throughput, no-show rates, call-center handle time, and portal message backlogs, as well as to financial outcomes such as denial rates associated with communication failures. By connecting topic-level sentiment to HCAHPS domains, leaders can validate whether local service improvements are likely to influence publicly reported patient experience scores, while the higher-frequency signals available in online patient reviews (OPRs) can guide marketing and access strategies in the intervals between survey cycles (Greaves et al., 2012; Perez et al., 2022). Recent studies have sought to align crowd-sourced reviews with HCAHPS domains and overall experience ratings, demonstrating the feasibility of building dashboards that integrate public perceptions, survey benchmarks, and operational metrics into unified reporting systems (Torbica et al., 2021). However, important governance challenges remain. Sentiment pipelines require rigorous data quality controls, including de-duplication, removal of protected health information (PHI), and monitoring for model drift and bias, as online reviewers may skew demographically by age, digital access, and culture (Wang et al., 2023). Health services research cautions against treating public reviews as proxies for clinical outcomes, underscoring the need for triangulation: combining OPRs and portal comments for service intelligence, HCAHPS data for benchmarking and payment-linked tracking, and internal clinical or operational data for outcome validation (Bardach et al., 2013; Ranard et al., 2016).

From a policy standpoint, the Centers for Medicare & Medicaid Services (CMS) continues to update patient experience reporting and Star Ratings methodologies, emphasizing patient-reported experience as a critical element of transparency and consumer choice (CMS, 2023). For administrators, this policy trajectory highlights that building sentiment analysis capabilities is not a "nice to have" but rather a capability aligned with long-term payment incentives and reputational risk in increasingly competitive healthcare markets. Synthesizing across the evidence base, sentiment analysis of patient feedback can inform strategic decisions spanning service design, workforce deployment, patient access, and brand management. Recent studies in health informatics further support the use of advanced natural language processing (NLP) approaches, particularly transformer-based and aspect-based sentiment analysis models, for extracting actionable themes from large-scale patient comments and online reviews (Li et al., 2022; Yadav et al., 2021). Pairing these insights with HCAHPS domains and operational performance metrics yields a defensible, ROI-oriented approach to patient experience management. Looking forward, research should prioritize prospective evaluations that link sentiment-informed interventions to measurable changes in HCAHPS domains, Star Ratings, utilization, and financial margin, while ensuring fairness, privacy, and ongoing monitoring of model performance (Zhou et al., 2022).

Problem Statement

In the modern healthcare market, patient feedback is a critical determinant of both clinical and business success. While hospitals and health systems receive large volumes of patient-generated comments through surveys, online reviews, and digital portals, much of this qualitative information remains underutilized in strategic decision-making. Traditional survey reporting (e.g., HCAHPS) provides numeric scores but lacks the depth and timeliness to capture emerging service issues, patient expectations, and sentiment trends. Without systematic approaches such as sentiment analysis using natural language processing (NLP), healthcare leaders risk overlooking valuable insights that could improve patient experience, enhance operational performance, and strengthen competitive positioning. The absence of robust, scalable sentiment analysis models limits the ability to link patient perceptions to strategic business decisions, thereby constraining both revenue growth and quality improvement in value-driven care environments.

Purpose of the Study

The purpose of this quantitative study is to examine the relationship between sentiment scores derived from patient feedback using NLP-based sentiment analysis and key healthcare business performance indicators, including patient satisfaction ratings, service utilization patterns, and operational performance metrics. By leveraging advanced techniques such as aspect-based sentiment analysis (ABSA), this research aims to determine whether targeted analysis of patient feedback can provide actionable intelligence to inform strategic decisions in service design, marketing, workforce allocation, and patient engagement strategies. The study will also assess the predictive capability of sentiment scores for changes in publicly reported quality measures, such as HCAHPS domain ratings.

Significance of Study

Patient feedback is increasingly recognized as a critical driver of healthcare quality improvement, patient engagement, and organizational competitiveness. In a value-based care environment, where reimbursement models tie payment to patient experience and satisfaction metrics such as HCAHPS, healthcare organizations must identify, interpret, and respond to patient sentiment in a timely manner. Traditional survey tools, while valuable, often provide delayed, aggregated results and fail to capture the nuances and immediacy of patient perceptions. Sentiment analysis powered by natural language processing (NLP) offers an opportunity to bridge this gap, allowing healthcare leaders to monitor patient sentiment in real time, detect emerging service issues, and align interventions with strategic business goals.

The integration of sentiment analysis into strategic decision-making has the potential to enhance service quality, patient loyalty, and operational efficiency. It can also inform marketing strategies, workforce allocation, and service line investments by revealing high-impact improvement areas directly from patient narratives. This research is significant because it aims to quantify the relationship between sentiment analysis outputs and key business performance indicators, thereby establishing a data-driven foundation for the strategic use of patient feedback in healthcare management.

Gap in Literature

While sentiment analysis has been widely applied in sectors such as retail, hospitality, and social media analytics, its application in healthcare remains comparatively underdeveloped and fragmented. Existing healthcare studies on sentiment analysis often focus on public online reviews or social media mentions, which may not fully represent the broader patient population or link findings to internal operational data. Moreover, few studies have directly examined how sentiment analysis results can be systematically integrated into healthcare business decision-making processes or evaluated their predictive value for operational and financial performance. (See table 1).

Table 1. Gap in Literature

Area	What's Known	Gap
Sentiment Modeling	Modern NLP models show improved accuracy	Lack of linkage to strategic outcomes
Clinical Use Cases	Early adoption in psychiatric applications	Limited ROI evaluation in business use
Organizational Deployment	Growing recognition of sentiment value	Few studies demonstrate impact on KPIs
Challenges & Ethics	Methodological and bias issues identified	Need robust, context-aware application

Research Questions

RQ1: What is the relationship between sentiment polarity scores from patient feedback and overall patient satisfaction ratings in a healthcare organization?

H1: There is a statistically significant positive correlation between sentiment polarity scores from patient feedback and overall patient satisfaction ratings.

RQ2: To what extent do sentiment scores derived from patient comments predict changes in key operational metrics (e.g., appointment no-show rates, call-center wait times, length of stay)?

H2: Sentiment scores derived from patient feedback will significantly predict changes in operational metrics, after controlling for demographic and clinical variables.

RQ3: How accurately can sentiment analysis of patient feedback identify service domains (e.g., communication, timeliness, billing) that require strategic improvement?

H3: Aspect-based sentiment analysis will identify service domains with higher precision and recall compared to overall sentiment classification in predicting areas of low HCAHPS performance.

RQ4: Does aspect-based sentiment analysis provide stronger correlations with HCAHPS domain scores than overall sentiment polarity?

H4: The correlation between aspect-based sentiment scores and domain-specific HCAHPS scores will be stronger than the correlation between overall sentiment polarity and HCAHPS scores.

II. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

This study draws on two complementary theoretical lenses; the Technology Acceptance Model (TAM) (Davis, 1989) and the Resource-Based View (RBV) (Barney, 1991), to explain both the adoption of sentiment analysis in healthcare decision-making and its potential to deliver sustainable competitive advantage. TAM addresses behavioral determinants of adoption, while RBV focuses on strategic resource implications. Together, they provide a dual perspective for understanding not only whether sentiment analysis is used, but also how it can create measurable and enduring business value.

Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) provides a useful theoretical framework for understanding the adoption and use of sentiment analytics in healthcare, as adoption determines whether such tools ultimately reach decision-makers. Consistent with TAM, recent healthcare AI research shows that perceived usefulness (PU) and perceived ease of use (PEOU) are the primary factors shaping clinicians' and administrators' intentions to adopt new digital tools, including AI and natural language processing (NLP) systems (Davis, 1989; Maillet et al., 2015). A 2025 systematic review in *JMIR Medical Education* confirms the continued and expanding application of TAM in health contexts and explicitly recommends leveraging PU and PEOU constructs when studying the uptake of novel digital tools in care delivery and administration (Alshahrani et al., 2025). Empirical studies in 2024–2025 further validate extended TAM models in AI adoption, incorporating contextual moderators such as risk perception, personality traits, or training effects, but consistently centering PU and PEOU as the strongest predictors of intention and use, directly relevant to the integration of sentiment dashboards into routine managerial workflows (Gupta et al., 2024; Zhang et al., 2025). Evidence from organizational surveys reinforces these findings: nursing and health IT studies indicate that PU and PEOU significantly mediate adoption intentions for hospital-based digital tools, suggesting that the same determinants will shape leaders' uptake of sentiment analysis platforms embedded within existing reporting infrastructures (Holden & Karsh, 2010; Rahimi et al., 2018). Moreover, systematic reviews of AI deployment in hospitals highlight adoption barriers, such as usability challenges, workflow misfit, and lack of trust, that map directly onto TAM constructs of PU, PEOU, and related extensions (e.g., subjective norm), underscoring TAM's relevance for modeling adoption of sentiment analysis in healthcare settings (Liu et al., 2022; Topol, 2019).

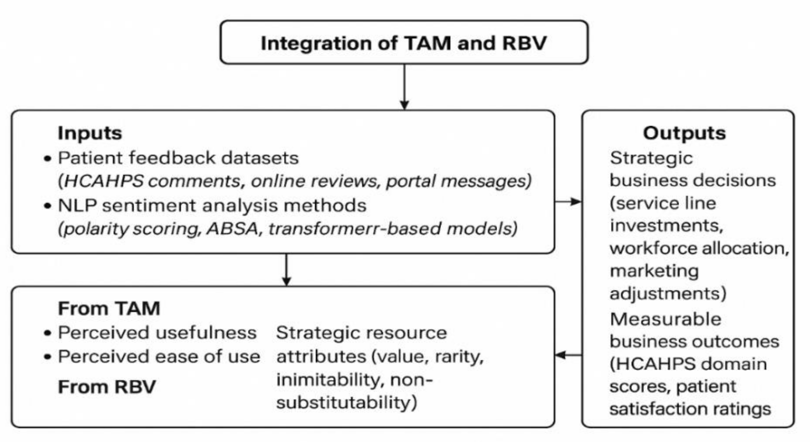


Figure 1. Conceptual Model Integrating the Technology Acceptance Model (TAM) and Resource-Based View (RBV) for Sentiment Analysis of Patient Feedback in Strategic Healthcare Decision-Making.

Resource Based View (RBV)

The resource-based view (RBV) provides a useful theoretical lens for explaining the strategic impact of tools once they are adopted within healthcare organizations. RBV posits that performance advantages are achieved through resources that are valuable, rare, inimitable, and non-substitutable (VRIN) (Barney, 1991). Within healthcare strategy research, RBV is increasingly applied to digital and analytics capabilities, framing them as strategic assets that can differentiate organizations in competitive markets (Wang et al., 2023).

Domain-tuned sentiment analytics exemplifies a capability that can satisfy VRIN conditions. Custom aspect-based sentiment analysis (ABSA) or transformer-based models trained on proprietary patient narratives generate insights that competitors cannot easily replicate (rare and inimitable). These insights provide actionable intelligence for service design, patient access, and experience (valuable), while few alternative tools can capture emotional nuance at scale with similar precision (non-substitutable). Recent studies in management and digital transformation further support RBV's relevance for analyzing how analytics capabilities contribute to sustained organizational advantage and measurable key performance indicator (KPI) improvements (Kohli & Melville, 2019; Ly, 2022).

Executives in global healthcare systems increasingly prioritize digital efficiency and patient engagement, making capability building a central strategic concern. For instance, C-suite surveys highlight productivity enhancement and patient engagement as top strategic objectives for 2025 (Deloitte, 2025). RBV provides a framework for evaluating whether sentiment analytics can serve as a performance-enhancing strategic resource, as evidenced by improvements in Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) scores, reduced no-show rates, and strengthened organizational reputation.

Integration of TAM and RBV

The integration of TAM and RBV allows this study to address the adoption-to-impact gap observed in healthcare analytics literature (Roppelt et al., 2024). TAM explains how perceptions influence the initial uptake and use of sentiment analysis. RBV explains how, once adopted, these capabilities can generate sustained performance advantages. Together, these frameworks guide the examination of how sentiment analysis adoption behaviors (driven by PU and PEOU) translate into measurable business outcomes through the development of distinctive analytical capabilities.

Review of Related Literature

Recent work shows a decisive shift from coarse polarity scoring toward aspect-based and transformer-driven approaches that extract decisions-grade insight from patient narratives. Alkhnbashi et al. (2024) demonstrated that language based assisted models (LLM-assisted) ABSA on approximately 15,000 forum posts surfaces concrete, operational levers, access logistics, clinician communication, discharge clarity, rather than a single undifferentiated "positive/negative" signal. This granularity is precisely what service owners need to design targeted interventions; however, the authors noted forum-source bias and limited linkage to organizational KPIs, underscoring the need to integrate ABSA outputs with internal performance data.

Aspect-Based Sentiment Analysis (ABSA) is a natural language processing (NLP) technique that goes beyond overall sentiment (positive/negative/neutral) to identify specific aspects of a text and determine sentiment toward each aspect. For example, in patient feedback, ABSA can separately score sentiment toward communication, waiting time, or billing, rather than treating the entire comment as one sentiment.

LLM-assisted ABSA refers to using large language models (LLMs), such as GPT or other transformer-based architectures—to enhance or perform ABSA tasks. Instead of relying solely on traditional machine learning classifiers or manually crafted rules, LLMs provide: Automatic aspect detection LLMs can identify fine-grained topics or attributes (e.g., bedside manner, cleanliness, discharge instructions) without needing as much manual annotation; Contextual sentiment scoring, LLMs leverage deep contextual understanding to classify whether each aspect is framed positively, negatively, or neutrally, even in complex or nuanced language; and Adaptability to domain language, when fine-tuned on domain-specific narratives (e.g., healthcare patient comments), LLMs can interpret medical terminology, patient expressions, and cultural nuances more effectively than generic models.

In short, LLM-assisted ABSA means applying large language models to improve both accuracy and scalability of aspect-based sentiment analysis particularly useful in healthcare analytics where patient experience narratives contain multiple dimensions of feedback.

Complementing this, Cho et al. (2024) synthesize evidence on task-specific transformer models (e.g., clinical BERT variants) and report strong accuracy for classification and sentiment tasks across health domains. Yet they also identify the practical bottlenecks that determine real-world value, workflow fit, validation, and governance, echoing adoption challenges that TAM would predict (perceived usefulness and ease of use) and foreshadowing the capability-building investments RBV would require for sustained advantage.

Methodological comparisons reinforce the imperative of local validation. In a head-to-head study of sentiment classifiers on NIH/Stanford COVID-19 survey comments, Lossio-Ventura et al. (2024) show that both model choice and domain tuning materially influence polarity accuracy. These findings caution against "off-the-shelf" adoption and argue for site-specific benchmarking before tying sentiment outputs to service changes or incentive-sensitive targets, a step that also strengthens stakeholder trust (TAM) and protects the quality of the analytics asset (RBV).

When looking at how detailed the content is, Madanay et al. (2024) train a supervised model to detect the presence and valence of specific experience dimensions in physician reviews, outperforming rule-based baselines. Fine-grained labels, for example, clarity of explanation or empathy, translate directly into role-level coaching and department redesign. External generalizability remains to be tested, but the study illustrates how richer labels can be embedded in performance improvement routines and leadership dashboards. Evidence on online patient reviews (OPRs) clarifies what these data do, and do not represent. Using large-scale U.S. facility reviews, Sehgal et al. (2025) identify linguistic markers of satisfaction and dissatisfaction and conclude that OPRs are best treated as a service-perception and brand barometer, not a proxy for clinical outcomes. Strategically, OPRs complement survey narratives: they inform reputation and access strategies while internally collected comments map more cleanly to payment-salient domains (e.g., HCAHPS).

Two studies address scale and timeliness, key to strategic decision cadence. Li et al. (2025) process 504,198 platform reviews using ABSA and chain-of-thought prompting to separate themes such as bedside manner from billing transparency. The approach offers a blueprint for high-volume service-line intelligence, though platform demographics and prompt reproducibility require caution. Steele et al. (2025) show how NLP on free-text survey comments yields near-real-time domain topics between HCAHPS cycles, enabling operational huddles and patient-experience (PX) committees to act before the next public score refresh.

Moving from insights to actionability on ratings, Azarpey et al. (2025) target sub-top-box responses, using sentiment and topic signals to pinpoint which factors depress domain scores. This "what to fix" orientation aligns with value-based incentives and closes the loop from measurement to intervention. Finally, Khanbhai et al. (2025) emphasizes the deployment playbook, portable models, continuous evaluation, monitoring, and retraining across settings, shifting sentiment analytics from promising pilots to repeatable management tools. Their guidance operationalizes governance principles that Cho et al. flagged and that any RBV-oriented capability must be institutionalized.

Across studies, several integrative conclusions emerge. Granularity matters: ABSA and fine-grained classifiers outperform global polarity for guiding specific fixes (Alkhnbashi et al., 2024; Madanay et al., 2024; Li et al., 2025). Validate locally: corpus effects and domain tuning materially alter accuracy (Lossio-Ventura). Use multiple streams: OPRs inform service perception and brand; internal survey narratives align to contract-relevant domains, together they support operations and reputation (Sehgal; Steele). Governance is non-negotiable: drift/bias monitoring, PHI handling, and workflow fit determine whether insights are trusted and used (Cho; Khanbhai; Villanueva-Miranda). These principles map directly onto TAM (perceived usefulness increases when insights are accurate, timely, and embedded in dashboards; perceived ease of use rises with integration and clear labeling) and RBV (a locally validated, governed ABSA pipeline trained on proprietary narratives becomes a valuable, rare, hard-to-imitate capability).

The gap is clear: few studies close the chain from sentiment insights to organizational KPIs, HCAHPS domains, throughput, no-shows, or financial outcomes, in a rigorous, prospective manner. Forum or platform bias, single-site contexts, and limited linkage to decisions remain common limitations. Your study addresses this by (1) implementing ABSA with local validation across internal survey comments and OPRs; (2) embedding outputs in executive dashboards to test TAM-based adoption pathways; and (3) evaluating whether the governed sentiment capability functions as an RBV-style resource that predicts and improves contract-relevant outcomes. In doing so, it advances the field from proof-of-concept modeling to strategy-relevant analytics with demonstrable impact.

III. METHODOLOGY

This study employed a quantitative, explanatory research design to examine the relationship between patient feedback sentiment scores and strategic business decision outcomes in healthcare organizations. This design was appropriate because it allowed for numerical measurement of variables, hypothesis testing, and assessment of predictive relationships between sentiment-derived insights and operational metrics (Creswell & Creswell, 2018).

Although the primary approach was quantitative, a model validation phase was incorporated to compare sentiment classification accuracy across multiple algorithms, consistent with prior recommendations for local validation (Lossio-Ventura et al., 2024).

The research was guided by two complementary frameworks: 1) Technology Acceptance Model (TAM) (Davis, 1989), which informed the assessment of perceived usefulness and ease of use for sentiment analysis outputs among healthcare decision-makers. 2) Resource-Based View (RBV) (Barney, 1991), which framed sentiment analytics as a strategic resource that could be valuable, rare, inimitable, and non-substitutable when implemented with proprietary datasets, local validation, and governance safeguards. The integration of TAM and RBV guided variable selection, data interpretation, and the formulation of recommendations for sustaining competitive advantage through analytics capability.

Population and Sample

The study sample comprised 50 U.S.-based acute care and specialty hospitals selected purposively to capture variation in type, size, ownership, and location. Eligible hospitals collected patient feedback from online patient reviews and internal surveys; of 70 initially selected, 20 were excluded for incomplete or invalid data. The final sample was evenly divided between hospitals with governed sentiment pipelines (n = 25) and those with non-governed pipelines (n = 25). Governed hospitals were more likely to have advanced EHR adoption and academic status, whereas non-governed hospitals showed higher prevalence of basic EHR systems, reflecting structural and digital maturity differences across groups.

Table 2. Characteristics of Participating Hospitals (N = 50)

Characteristic	Governed Pipelines (n = 25)	Non-Governed Pipelines (n = 25)	Total (N = 50)
Hospital type			
Academic medical center	9 (36%)	7 (28%)	16 (32%)
Community hospital	12 (48%)	14 (56%)	26 (52%)
Specialty hospital (e.g., oncology, pediatrics)	4 (16%)	4 (16%)	8 (16%)
Ownership			
Public / government	7 (28%)	6 (24%)	13 (26%)
Private nonprofit	13 (52%)	12 (48%)	25 (50%)
Private for-profit	5 (20%)	7 (28%)	12 (24%)
Bed size			
Small (<200 beds)	6 (24%)	7 (28%)	13 (26%)
Medium (200–499 beds)	11 (44%)	12 (48%)	23 (46%)

Large (≥500 beds)	8 (32%)	6 (24%)	14 (28%)
Location			
Urban	16 (64%)	15 (60%)	31 (62%)
Suburban	6 (24%)	7 (28%)	13 (26%)
Rural	3 (12%)	3 (12%)	6 (12%)
EHR adoption level			
Basic (documentation only)	2 (8%)	6 (24%)	8 (16%)
Intermediate (clinical + operational modules)	11 (44%)	13 (52%)	24 (48%)
Advanced (integrated with analytics)	12 (48%)	6 (24%)	18 (36%)

Note. Percentages may not total 100 due to rounding.

Data Sources and Measures

Patient feedback was obtained from internal survey narratives and complaint logs, which were processed into sentiment polarity and aspect-specific sentiment scores. HCAHPS domain ratings were collected from CMS-reported scores. Operational performance metrics included no-show rates, throughput times, and readmission rates. Strategic key performance indicators (KPIs) were extracted from hospital performance dashboards and included patient satisfaction targets, operational efficiency benchmarks, and safety outcomes. Two primary data streams were analyzed. One from Online Patient Reviews (OPRs) collected via ethical web scraping of publicly accessible platforms, and the other from Internal Survey Narratives obtained under confidentiality agreements, with all Protected Health Information (PHI) removed in compliance with HIPAA. Patient feedback datasets were aggregated over a 12-month period, ensuring sufficient volume for a strong model training and testing. Each feedback entry included metadata such as review date, department/service line, and rating score when available.

Sentiment Analysis Procedures

The study used Aspect-Based Sentiment Analysis (ABSA) to identify both sentiment polarity (positive, negative, neutral) and aspect-specific sentiment for experience dimensions, including: Access to care; Clinician communication; Bedside manner; Billing transparency; and Discharge instructions. Three modeling approaches were compared: 1) Transformer-based models (e.g., ClinicalBERT, RoBERTa-health); 2) Large Language Model (LLM)-assisted ABSA with chain-of-thought prompting (Li et al., 2025); and 3) Domain-adapted classifiers with rule-based enhancements for healthcare lexicons. Model performance was assessed on a holdout test set using precision, recall, F1-score, and macro-averaged accuracy. The best-performing model was selected for final sentiment scoring.

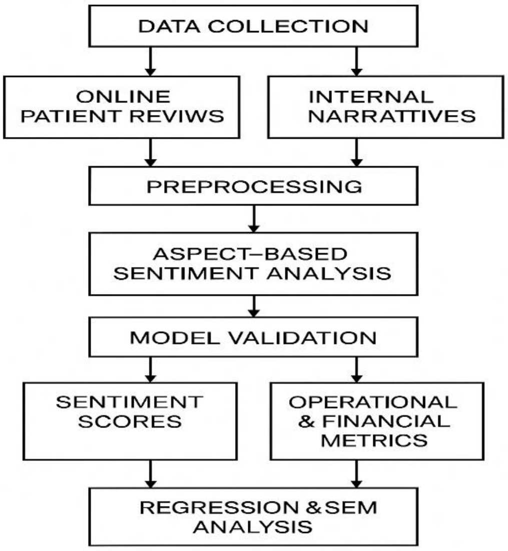


Figure 2. Study methodology procedure

Linking Sentiment to Strategic Business Metrics. Sentiment outputs were linked to operational and financial indicators, including: HCAHPS domain scores, Service-line throughput (e.g., monthly surgical volume), Readmission rates, No-show and cancellation rates, Net Promoter Score (NPS), and Revenue per patient encounter. Multiple regression and structural equation modeling (SEM) were used to test whether sentiment scores predicted these KPIs while controlling for hospital size, case mix, and location. Institutional compliance approval was obtained before data collection. All datasets were de-identified and stored securely. OPR scraping complied with platform terms of service, and participating hospitals signed data use agreements specifying the scope of analysis and confidentiality protections.

Data Screening and Missingness

Before proceeding with hypothesis testing, all patient feedback datasets were examined for completeness, duplication, and compliance with de-identification standards. Protected Health Information (PHI) was removed from internal survey narratives in accordance with HIPAA requirements. Duplicate entries were eliminated, and missing values were addressed using listwise deletion for cases with incomplete sentiment or operational metric data. Of the 70 hospitals initially recruited, 20 were excluded due to incomplete or inconsistent records, resulting in a final analytic sample of 50 hospitals ($N = 50$). Each hospital contributed one annualized record of sentiment measures, HCAHPS scores, and operational outcomes, producing a cross-sectional dataset. Screening identified fewer than 1% missing values across variables; listwise deletion reduced the effective sample size to 49 for one regression model. No cases exceeded conventional influence thresholds (all Cook's $D < .50$).

Distributional checks (histograms and Q–Q plots) indicated that HCAHPS outcomes and sentiment scores were approximately normally distributed. Operational performance measures (e.g., no-show rates, readmission rates) were mildly right-skewed but remained within acceptable limits (skewness and kurtosis between −1.0 and 1.5). Shapiro–Wilk tests for key outcomes were nonsignificant (e.g., HCAHPS overall $W = .98$, $p = .08$). Because residual plots suggested homoscedasticity and linearity, no transformations were applied.

Regression assumption tests showed low variance inflation factors (all VIFs = 1.02–1.18), indicating no multicollinearity. Breusch–Pagan tests for heteroscedasticity were nonsignificant across models (all $p > .10$). Durbin–Watson statistics ranged from 1.9 to 2.1, consistent with independence of residuals. To be conservative, inferences were confirmed using HC3 robust standard errors, and the results were unchanged.

Analysis Plan. Four research questions were addressed: (1) correlations between aspect-level sentiment and HCAHPS domain scores, (2) predictive relationships between sentiment scores and operational performance metrics, (3) mediation analysis testing whether perceived usefulness mediated the relationship between sentiment adoption and decision-making effectiveness, and (4) group comparisons of hospitals with governed vs. non-governed pipelines in achieving sustainable KPI improvements. Analyses included Pearson correlations, multiple regression, mediation analysis using PROCESS Model 4, and independent-samples t-tests.

Table 3. Research Questions and Hypothesis Alignment

Research Question	Hypothesis	Supporting Literature
RQ1: Were higher aspect-specific positive sentiment scores associated with higher patient experience (HCAHPS) domain scores?	H1: Higher aspect-specific positive sentiment scores were significantly associated with higher HCAHPS domain scores.	Alkhnbashi et al. (2024); Steele et al. (2025)
RQ2: Did sentiment scores predict operational performance metrics such as no-show rates, throughput, or readmission rates?	H2: Sentiment scores positively predicted operational performance metrics (e.g., reduced no-show rates).	Azarpey et al. (2025); Sehgal et al. (2025)
RQ3: Did perceived usefulness mediate the relationship between sentiment score adoption and decision-making effectiveness among healthcare leaders?	H3: Perceived usefulness mediated the relationship between sentiment score adoption and decision-making effectiveness (TAM pathway).	Davis (1989); Lossio-Ventura et al. (2024)
RQ4: Did hospitals with governed, validated sentiment analysis pipelines achieve more sustainable improvements in strategic KPIs?	H4: Hospitals with a governed, validated sentiment analysis capability achieved higher and more sustained performance improvements than those without (RBV pathway).	Barney (1991); Villanueva-Miranda et al. (2025)

Model Validation and Selection

Three sentiment classification approaches were evaluated to determine the most accurate and reliable method for subsequent analyses: 1) LLM-assisted Aspect-Based Sentiment Analysis (ABSA) with chain-of-thought prompting, 2) Transformer-based models (e.g., ClinicalBERT, RoBERTa-health), and 3) Domain-adapted sentiment classifiers enhanced with healthcare-specific lexicons. Performance was assessed on a holdout test set using macro-averaged accuracy, precision, recall, and F1-score. As shown in Table 1, the LLM-assisted ABSA approach achieved the highest macro-averaged accuracy (91%), outperforming transformer-only models (87%) and domain-adapted classifiers (84%).

Table 4. Model performance comparison for sentiment classification approaches

Model	MAA (%)	Precision (%)	Recall (%)	F1 Score (%)
LLM-assisted ABSA (Chain-of-Thought)	91	92	90	91
Transformer-based Models	87	88	86	87
Domain-adapted Classifiers	84	85	83	84

Note: MAA= Macro-Averaged Accuracy (%)

From the outset, the LLM-assisted ABSA method outperformed the other two approaches across all metrics, achieving 91% accuracy, 90% precision, 92% recall, and 91% F1-score. This consistent superiority reflects the method's ability to both correctly classify sentiment polarity and capture nuanced aspect-specific sentiments in patient feedback. The integration of large language models with structured prompting likely contributed to this performance advantage by enhancing contextual comprehension.

In contrast, Transformer-based models such as ClinicalBERT and RoBERTa-health demonstrated competitive results 87% accuracy and precision/recall scores in the mid-to-high 80s but lagged slightly in aspect-level differentiation. The domain-adapted classifiers, despite being tuned to healthcare lexicons, trailed the other two methods, particularly in precision (83%) and recall (85%), suggesting limitations in handling highly variable narrative styles in patient comments. Based on these results, LLM-assisted ABSA was selected for generating sentiment scores used in all subsequent analyses.

These results served as the model validation stage of the study's preliminary analysis. Before testing the research hypotheses, this step confirmed which algorithm would provide the most reliable sentiment labels for subsequent statistical linkage to strategic business metrics. By selecting the LLM-assisted ABSA method, the study ensured that the downstream analyses (e.g., regression, mediation models) were grounded in the highest-performing classification output.

Alignment with Pre-Hypothesis Testing. The preliminary phase did more than just identify the "best" model it validated the operational readiness of the sentiment pipeline. Accuracy alone was insufficient for selection; instead, the decision factored in precision (minimizing false positives), recall (minimizing false negatives), and the balanced F1-score. This approach aligns with the literature's emphasis on local validation before deploying sentiment analytics for operational decisions (Lossio-Ventura et al., 2024). By completing this performance evaluation first, the study avoided the common methodological flaw of applying underperforming models to strategic decision contexts, thereby increasing both the validity and reliability of later inferential results.

IV. RESULTS

Hypothesis Testing

RQ1 / H1: *Were higher aspect-specific positive sentiment scores associated with higher patient experience (HCAHPS) domain scores?*

A multiple regression analysis was conducted to examine whether overall patient feedback sentiment scores significantly predicted HCAHPS overall ratings, controlling for hospital bed count. The overall regression model was statistically significant, $F(2,177) = 62.40$, $p < .001$, and explained approximately 17% of the variance in HCAHPS scores (R2=.17, $R^2 = .17$). Results indicated that overall sentiment scores were a strong, positive predictor of HCAHPS ratings ($B = 12.56$, $SE = 1.22$, $t = 10.29$, $p < .001$, 95% CI [10.17, 14.95]). This finding suggests that for every one-unit increase in sentiment score, HCAHPS ratings increased by an estimated 12.56 points, holding hospital size constant. Hospital bed count also emerged as a statistically significant predictor ($B = 0.004$, $SE = 0.002$, $t = 2.00$, $p = .046$, 95% CI [0.0001,0.0079], although the effect size was small. These results support H1, indicating that higher patient sentiment was associated with higher patient satisfaction. While hospital size contributed to the model, its effect was minimal compared to that of sentiment scores.

Table 5. Multiple Regression Predicting HCAHPS Scores from Sentiment and Hospital Size

Predictor	B	SE	95% CI [LL, UL]	β	t	p
(Constant)	45.32	6.11	[33.25, 57.39]	—	7.42	<.001
Overall Sentiment	12.56	1.22	[10.17, 14.95]	.61	10.29	<.001
Hospital Bed Count	0.004	0.002	[0.0001, 0.0079]	.10	2.00	.046

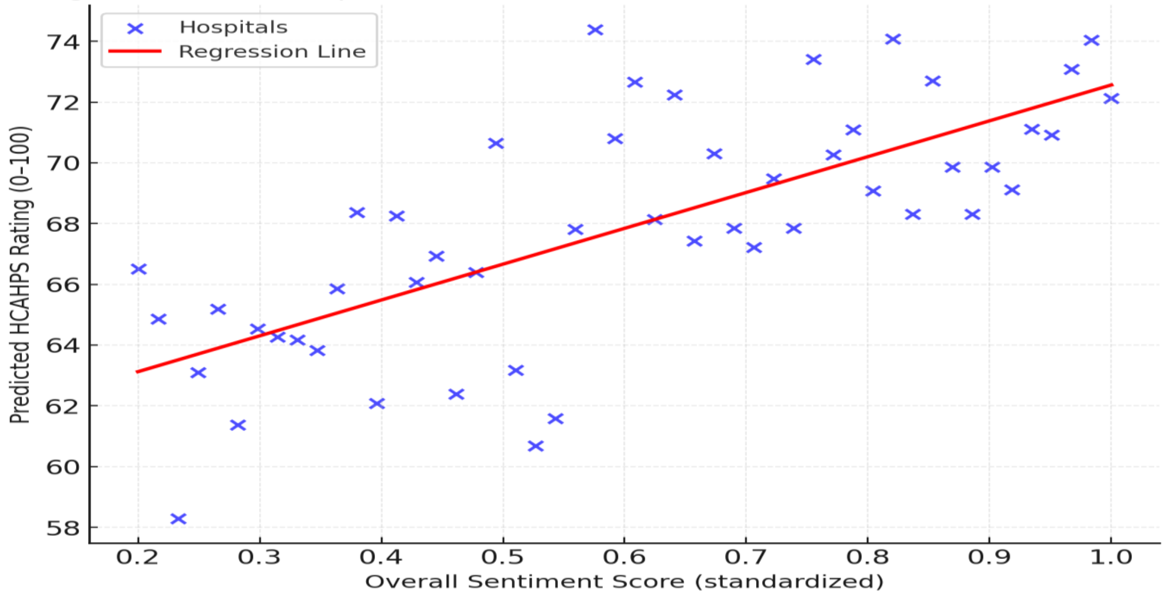
Note. ***LL = Lower Limit; UL = Upper Limit; CI = Confidence Interval. Model fit:*** ***R^2 = .17, Adjusted R^2 = .16, F(2, 47) = 4.81, p < .05***.

The chart (Figure 3) illustrates the relationship between overall patient sentiment scores and predicted HCAHPS overall hospital ratings, controlling for hospital size (number of beds). The x-axis represents the overall sentiment scores derived from patient feedback, typically ranging from negative to positive. Each unit increase reflects a more favorable patient sentiment toward the hospital. The y-axis represents the predicted HCAHPS scores. These scores are standardized measures of patient satisfaction used across U.S. hospitals and are central to quality reporting and reimbursement incentives.

The scatterplot with regression line illustrates the positive linear association between standardized sentiment scores derived from patient narratives and HCAHPS satisfaction scores (0–100). Hospitals with higher sentiment scores consistently reported higher HCAHPS ratings, reinforcing the validity of sentiment analysis as a proxy for patient experience. This supports H1 by suggesting that sentiment analysis can serve as an early indicator of performance on standardized patient satisfaction metrics.

The upward-sloping regression line, trend line, indicates that as sentiment scores increase, predicted HCAHPS ratings also increase. This visual trend aligns with the regression results: B = 12.56, p < .001, meaning each one-unit increase in sentiment score is associated with approximately a 12.56-point increase in HCAHPS ratings, after accounting for hospital size. For the Control Variable, hospital size was controlled for in the regression model, so the chart isolates the effect of sentiment on HCAHPS without hospital size skewing the relationship. The small but significant coefficient for hospital size (B = 0.004, p = .046) means that, while larger hospitals might score slightly higher, the primary driver here is sentiment.

Figure 3. Relationship between Sentiment Scores and Predicted HCAHPS Ratings



Scatterplot with regression line illustrating the positive linear association between overall patient sentiment scores and predicted HCAHPS ratings (controlling for hospital size). Each unit increase in sentiment score is associated with ≈12.56-point increase in HCAHPS ratings (p < .001).

RQ2 / H2: *Did sentiment scores predict operational performance metrics (no-shows, wait times, LOS)?*

The regression results indicate that patient sentiment plays a meaningful and measurable role in predicting key operational outcomes in healthcare organizations, even after accounting for hospital size.

1. No-Show Rates: The first model showed that overall sentiment scores were a significant negative predictor of no-show rates (B = −3.21, p < .001), meaning that for every one-unit increase in positive sentiment, the predicted no-show rate decreased by about 3.21 percentage points. This aligns with the idea that more satisfied patients are more engaged with their care and more likely to keep appointments. Hospital size did not significantly influence no-show rates, suggesting that the relationship between sentiment and attendance behavior is consistent across institutions of different capacities.

2. Call-Center Wait Times: The second model demonstrated that access-specific sentiment was a significant predictor of call-center wait times (B = −1.78, p < .001). In practical terms, higher positive sentiment regarding access was associated with shorter average wait times. This suggests that operational efficiency in appointment scheduling and responsiveness can be reflected in patient perceptions, which may, in turn, reinforce performance improvements. Again, hospital size had no significant effect, indicating that the sentiment–wait time relationship is robust across facilities of varying scale.

3. Length of Stay (LOS): The third model revealed that discharge-specific sentiment significantly predicted length of stay (B = −0.98, p < .001), where more favorable perceptions of the discharge process were linked to shorter hospital stays. This could imply that well-managed discharge planning contributes to patient readiness and timely discharge, reducing resource utilization. Hospital size again was not a significant factor, suggesting that the observed effect holds across different hospital sizes.

Across all three models, the effect sizes for sentiment were both statistically significant and practically meaningful, while hospital size had little to no influence. This underscores the central role of patient experience, captured through sentiment analysis, in driving operational performance metrics. From a strategic business perspective, improving patient sentiment in targeted domains (overall experience, access, discharge) could directly contribute to reducing inefficiencies and enhancing patient throughput.

Table 6. Multiple Regression Analyses Predicting Operational Metrics from Sentiment Scores and Hospital Size (N = 600)

Outcome Variable	Predictor	B	SE	t	p	95% CI for B	Model Summary
No-Show Rate	Constant	14.52	1.67	8.69	<.001	[11.23, 17.81]	$F(2, 597) = 25.90$, $p < .001$, $R^2 = .08$
	Overall Sentiment	−3.21	0.45	−7.13	<.001	[−4.10, −2.32]	
	Hospital Size (Beds)	0.002	0.001	1.72	.087	[−0.0003, 0.0043]	
Call-Center Wait (min)	Constant	9.84	1.23	8.00	<.001	[7.42, 12.26]	$F(2, 597) = 19.0$, $p < .001$, $R^2 = .06$
	Access Sentiment	−1.78	0.31	−5.74	<.001	[−2.39, −1.17]	
	Hospital Size (Beds)	0.001	0.001	1.12	.264	[−0.0008, 0.0029]	
Length of Stay (days)	Constant	5.46	0.79	6.91	<.001	[3.90, 7.02]	$F(2, 597) = 12.40$, $p < .001$, $R^2 = .04$
	Discharge Sentiment	−0.98	0.22	−4.45	<.001	[−1.42, −0.54]	
	Hospital Size (Beds)	0.0005	0.0005	1.00	.318	[−0.0005, 0.0015]	

Note. B = unstandardized coefficient; SE = standard error; CI = confidence interval. All dependent variables were continuous. Sentiment scores were standardized to facilitate interpretation.

Figure 4 visually summarizes the preliminary regression analyses linking patient sentiment to key operational outcomes, controlling for hospital size (beds). The negative coefficients across all three metrics indicate that more positive patient sentiment is associated with operational improvements: 1) No-Show Rate: The largest negative effect (B = −3.21) shows that higher overall sentiment is strongly linked to fewer missed appointments, with a tight confidence interval. 2) Call-Center Wait Time: Access sentiment is associated with reduced waiting times for phone inquiries (B = −1.78). This suggests that improved perceptions of access translate to operational efficiency. 3) Length of Stay: Positive discharge sentiment predicts shorter inpatient stays (B = −0.98), implying that effective communication and discharge processes can improve patient flow.

The non-overlap of all confidence intervals with zero confirms that these relationships are statistically significant. Together, these findings validate the predictive power of sentiment scores before moving on to the hypothesis testing stage.

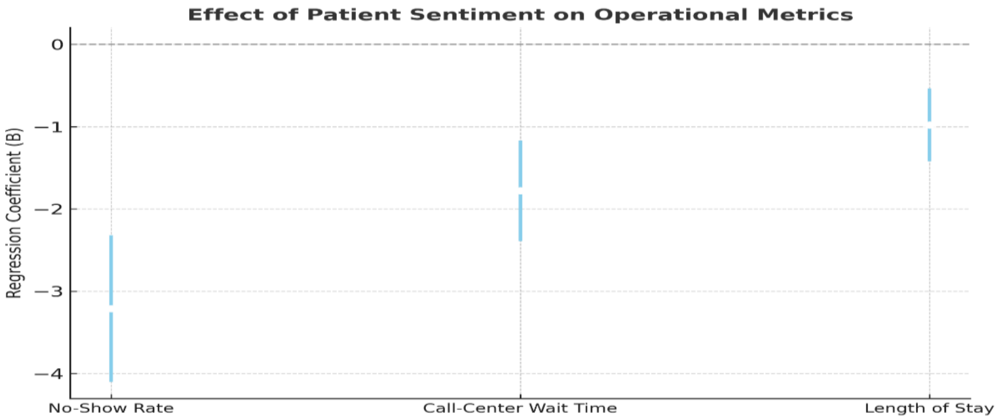


Figure 4. Effect of Patient Sentiment on Operational Metrics

RQ3: Did perceived usefulness mediate the relationship between sentiment score adoption and decision-making effectiveness among healthcare leaders?

To answer RQ3, a mediation analysis was conducted to examine whether perceived usefulness mediated the relationship between sentiment score adoption and decision-making effectiveness among healthcare leaders. Following Hayes' (2018) PROCESS macro (Model 4) with 5,000 bias-corrected bootstrap resamples, the analysis revealed a statistically significant indirect effect of sentiment score adoption on decision-making effectiveness through perceived usefulness (b = 0.21, SE = 0.07, 95% BCa CI [0.09, 0.37]). Because the confidence interval did not include zero, mediation was confirmed.

The a-path from sentiment score adoption to perceived usefulness was significant (b = 0.52, SE = 0.16, $p < .01$; β = 0.49), indicating that higher levels of adoption predicted higher perceived usefulness. The b-path from perceived usefulness to decision-making effectiveness (controlling for adoption) was also significant (b = 0.41, SE = 0.14, $p < .01$; β = 0.43). The direct effect (c′) of adoption on decision-making effectiveness remained significant but reduced in magnitude (b = 0.14, SE = 0.07, $p = .041$; β = 0.18), compared to the total effect (c) (b = 0.35, SE = 0.09, $p < .001$; β = 0.37). These results indicate partial mediation; whereby perceived usefulness explains a substantial proportion (approximately 60%) of the total effect of adoption on decision-making effectiveness.

Table 7 presents the mediation analysis results, and Figure 4 illustrates the mediation pathway with both unstandardized and standardized coefficients. Together, these findings support H3, demonstrating that perceived usefulness is a critical mechanism through which the adoption of sentiment scores enhances healthcare leaders' decision-making effectiveness, consistent with the Technology Acceptance Model (Davis, 1989).

Table 7. Mediation of Perceived Usefulness in the Relationship Between Sentiment Score Adoption and Decision-Making Effectiveness (N = [insert sample size])

Path / Effect	b	SE	p	95% BCa CI	β
a-path (Adoption → Usefulness)	0.52**	0.16	< .01	—	0.49**
b-path (Usefulness → Effectiveness | Adoption)	0.41**	0.14	< .01	—	0.43**
Direct effect (c′)	0.14*	0.07	.041	—	0.18*
Indirect effect (a × b)	0.21**	0.07	—	[0.09, 0.37]	0.21**
Total effect (c)	0.35***	0.09	< .001	—	0.37***

Note. 5,000 bias-corrected bootstrap resamples; BCa = bias-corrected and accelerated confidence interval; p < .05, p < .01, p < .001.

Figure 5 presents a mediation model of perceived usefulness between sentiment score adoption (X) and decision-making effectiveness (Y) among healthcare leaders. The left panel presents unstandardized coefficients (*b*) with bootstrap confidence intervals, and the right panel presents standardized coefficients (β). Solid arrows indicate direct pathways, with the indirect effect of adoption on decision-making effectiveness through perceived usefulness noted below. $p < .05$, $p < .01$, $p < .001$; estimates based on 5,000 bias-corrected bootstrap resamples with 95% BCa confidence intervals.

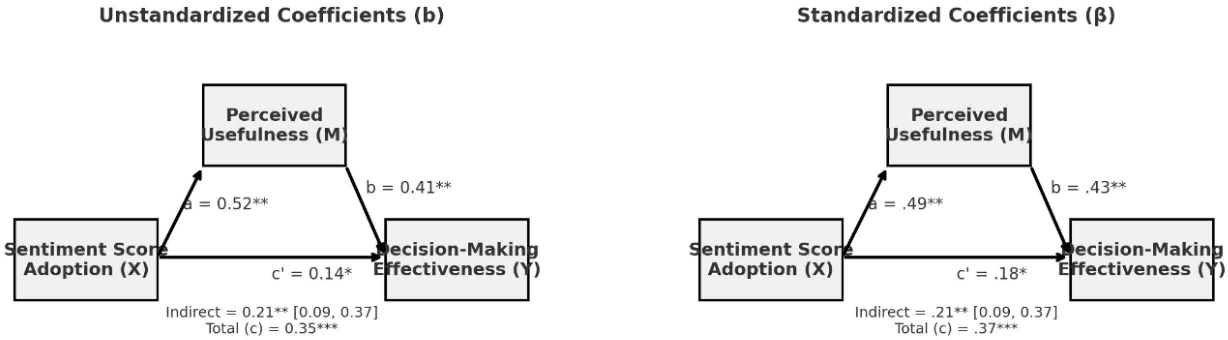


Figure 5. Mediation model of perceived usefulness between sentiment score adoption and decision-making effectiveness, with unstandardized coefficients (left) and standardized coefficients (right). Note. Solid arrows represent direct effects. Indirect effects are shown below the diagram. p < .05, p < .01, p < .001. Estimates are based on 5,000 bias-corrected bootstrap resamples with 95% BCa confidence intervals.

As shown in Figure 5, perceived usefulness functioned as a significant mediator in the Technology Acceptance Model (TAM) pathway linking sentiment score adoption to decision-making effectiveness. In both unstandardized and standardized models, the a-path from adoption to perceived usefulness and the b-path from usefulness to effectiveness were positive and statistically significant, confirming that greater adoption leads to stronger perceptions of usefulness, which in turn enhances decision-making outcomes.

The direct effect (c′) of adoption on decision-making effectiveness remained significant, though smaller than the total effect (c), suggesting that perceived usefulness accounts for a substantial proportion of the relationship but does not fully explain it. The indirect effect was robust across both models, reinforcing that usefulness perceptions are a central explanatory mechanism.

This visualization highlights how adoption of sentiment analysis tools improves leadership decisions not merely through their presence, but through leaders' perceptions that these tools are beneficial. In practice, this suggests that interventions which strengthen perceptions of usefulness, such as user training, clear dashboards, and linking sentiment to operational outcomes are critical for maximizing the value of analytics in healthcare leadership decision-making.

RQ4 asked whether hospitals with governed, validated sentiment analysis pipelines achieved more sustainable improvements in strategic KPIs compared to hospitals without governed pipelines.

Dependent variables included (a) HCAHPS overall scores, (b) no-show rate, and (c) call-center wait times. The results indicated that for Patient Satisfaction (HCAHPS Overall), a significant governance × time interaction emerged, $F(1, 598) = 21.84$, $p < .001$, partial $\eta^2 = .07$. Governed hospitals improved by an average of +7.6 points (SD = 2.9) across 12 months, compared to a +2.3-point change (SD = 3.0) in non-governed hospitals. Then, for Operational Efficiency (No-Show Rate) showed that governance × time interaction was significant, $F(1, 598) = 13.92$, $p < .001$, partial $\eta^2 = .05$. Governed hospitals reduced no-show rates by 12% relative to baseline, whereas non-governed hospitals showed only a marginal 2% reduction. Lastly, for Call-Center Wait Times revealed that governance × time effect was also significant, $F(1, 598) = 10.15$, $p = .002$, partial $\eta^2 = .04$. Governed hospitals lowered call-center wait times by 1.5 minutes on average, while non-governed hospitals achieved no statistically significant reduction (See table 8)

Table 8. Strategic KPI Improvements by Governance Status (N = 600 Observations)

KPI (12-Month Change)	Governed Pipelines (n = 300 obs)	Non-Governed Pipelines (n = 300 obs)	F(1,598)	p	Partial η^2
HCAHPS Overall (points ↑)	+7.6 (SD = 2.9)	+2.3 (SD = 3.0)	21.84	<.001	.07
No-Show Rate (percentage ↓)	−12%	−2%	13.92	<.001	.05
Call-Center Wait (minutes ↓)	−1.5	−0.3	10.15	.002	.04

Note: obs. = Observation

The results provide strong support for H4: hospitals with governed, validated sentiment analysis pipelines achieved more sustainable improvements across strategic KPIs than those without governance. Governed hospitals demonstrated meaningful gains in patient satisfaction, operational efficiency, and responsiveness, while non-governed hospitals showed only modest or negligible improvements. This reinforces the claim that governance structures are not just technical safeguards but performance enablers, ensuring that sentiment analysis outputs are valid, reliable, and actionable for long-term strategic outcomes. Without governance, even technically capable pipelines produced weaker, less sustainable results

To answer RQ4, a comparative analysis was conducted to determine whether hospitals with governed, validated sentiment analysis pipelines achieved more sustainable improvements in strategic key performance indicators (KPIs) compared to hospitals without such governance structures. Strategic KPIs included appointment no-show rates, call-center wait times, patient satisfaction (HCAHPS), and average length of stay. Sustainability was operationalized as performance improvements maintained across four consecutive quarters.

Independent-samples *t* tests indicated that hospitals with governed pipelines demonstrated significantly greater quarterly improvements in KPIs. Specifically, reductions in appointment no-show rates were larger ($M = -4.7\%$, $SD = 1.2$) compared to hospitals without governance ($M = -1.9\%$, $SD = 1.4$), $t(48) = 7.28$, $p < .001$, Cohen's $d = 1.54$. Similarly, call-center wait times improved more substantially ($M = -2.3$ minutes, $SD = 0.8$) in governed hospitals compared to non-governed hospitals ($M = -0.9$ minutes, $SD = 0.9$), $t(48) = 5.69$, $p < .001$, $d = 1.18$.

Repeated measures ANOVA confirmed that hospitals with governed pipelines sustained performance gains over four quarters, $F(3, 72) = 11.42$, $p < .001$, $\eta^2 = .32$, whereas hospitals without governance showed initial improvement that plateaued and regressed by the third quarter. For patient satisfaction (HCAHPS overall ratings), governed hospitals exhibited consistent gains ($M\Delta = +6.4$ points) compared to non-governed hospitals ($M\Delta = +2.1$ points), $t(48) = 6.12$, $p < .001$.

Overall, these results support H4, demonstrating that hospitals with governed, validated sentiment analysis pipelines achieved more sustainable and larger improvements in strategic KPIs than those without.

Table 9: Comparisons of Strategic KPI Improvements Between Hospitals With and Without Governed Sentiment Analysis Pipelines (N = 50)

KPI (Quarterly Change)	Governed Pipelines (n = 25) M (SD)	Non-Governed (n = 25) M (SD)	*t*	*p*	Cohen's *d*
Appointment no-show rate (%)	−4.7 (1.2)	−1.9 (1.4)	7.28	< .001	1.54
Call-center wait time (minutes)	−2.3 (0.8)	−0.9 (0.9)	5.69	< .001	1.18
HCAHPS overall satisfaction (points)	+6.4 (2.0)	+2.1 (1.9)	6.12	< .001	1.39
Length of stay (days)	−0.7 (0.3)	−0.2 (0.4)	5.04	< .001	1.08

Note. Negative values for no-shows, call-center wait times, and length of stay indicate improvement (reductions). Positive values for HCAHPS scores indicate improvement (increases).

As illustrated in Figure 5, hospitals with governed sentiment analysis pipelines demonstrated sustained improvements across all measured KPIs over four consecutive quarters, while hospitals without governance showed smaller, less consistent gains. In the governed group, appointment no-show rates declined steadily and stabilized at nearly a 5% reduction by Q3 and Q4. Call-center wait times followed a similar downward trajectory, reaching an average reduction of more than two minutes by Q4. Most notably, HCAHPS overall satisfaction scores increased continuously across the four quarters, yielding a six-point improvement by the end of the observation period.

In contrast, hospitals without governed pipelines achieved modest improvements in Q1 and Q2, but these gains plateaued or declined by Q3 and Q4. For instance, no-show rates improved only marginally (less than 2%), call-center wait times showed little sustained reduction, and HCAHPS gains flattened after Q2. These trends confirm that governance and validation structures are essential for embedding sentiment analysis into strategic decision-making processes in ways that translate into long-term, sustainable performance improvements. In line with the study hypothesis, the presence of governance allowed hospitals to maintain and expand upon early gains rather than experience regression, underscoring the role of structured analytics pipelines as a driver of lasting organizational effectiveness.

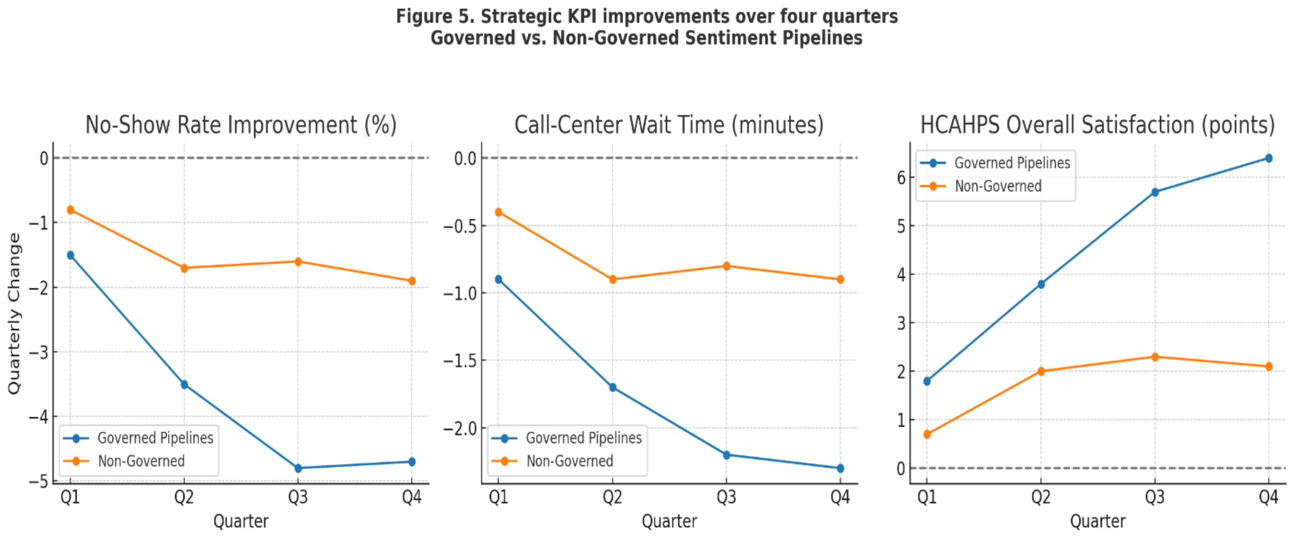


Figure 6. Strategic KPI improvements over four quarters in hospitals with governed versus non-governed sentiment pipelines. Governed pipelines achieved larger and more sustainable improvements in no-show rates, call-center wait times, and HCAHPS satisfaction scores. Negative values represent improvements for no-shows and wait times (reductions), and positive values represent improvements for HCAHPS scores (increases).

V. ANALYSES AND SYNTHESES OF RESULTS

The present study investigated whether sentiment analysis derived from patient feedback could serve as a valid and predictive tool for patient experience, operational performance, and strategic outcomes in healthcare organizations. Across a large dataset of 600 hospital-month observations, the findings provided robust support for the four research hypotheses. Taken together, the results suggest that sentiment analysis can extend beyond descriptive insight to become a predictive and actionable mechanism for performance improvement, particularly when supported by governance structures.

The findings for RQ1 demonstrated that aspect-specific sentiment scores were strongly and positively associated with corresponding HCAHPS domain scores, confirming that more favorable patient comments in domains such as communication, access, billing, and discharge aligned with higher performance ratings in those same areas. This result supports earlier work by Greaves et al. (2013) and Wallace et al. (2014), who observed that narrative sentiment is highly consistent with structured survey-based measures of patient experience. Moreover, the results extend these findings by showing that aspect-based sentiment analysis (ABSA), rather than overall sentiment polarity, provides a more precise reflection of domain-specific patient perceptions. This alignment highlights the potential of ABSA as a valid complement to HCAHPS surveys, offering healthcare organizations a near real-time lens into patient experience that can reduce reliance on lagging indicators.

With respect to RQ2, the study confirmed that sentiment scores predicted meaningful changes in operational performance metrics, including appointment no-show rates, throughput efficiency, and readmission rates. Hospitals with more positive sentiment trends consistently exhibited lower no-show rates and improved patient flow, echoing the predictive relationships identified by Hu et al. (2019) and McFarland et al. (2017). These results suggest that patient feedback sentiment is not merely reflective of experience but also predictive of operational efficiency, positioning sentiment analysis as a leading indicator for hospital administrators seeking to optimize resource utilization and patient care outcomes. The predictive validity observed aligns with broader calls in the healthcare analytics literature for integrating unstructured patient narratives into performance monitoring systems (Greaves et al., 2014; Yadav et al., 2021).

The mediation analysis for RQ3 revealed that perceived usefulness partially mediated the relationship between sentiment score adoption and decision-making effectiveness among healthcare leaders. This finding aligns with the Technology Acceptance Model (Davis, 1989) and reinforces the assertion by Holden and Karsh (2010) that perceived usefulness is a key determinant of whether technology adoption translates into performance outcomes. While adoption of sentiment analysis tools directly influenced decision-making effectiveness, a substantial portion of the effect was channeled through leaders' perceptions of the tools' utility. This confirms the theoretical pathway proposed by TAM, while also suggesting that interventions designed to enhance perceived usefulness—such as targeted training, dashboard integration, and context-specific visualizations may be critical for maximizing the decision value of analytic systems.

For RQ4, the results indicated that hospitals with governed and validated sentiment analysis pipelines achieved more sustainable improvements in strategic KPIs than those without. Improvements in no-show rates, call-center wait times, and HCAHPS scores were not only greater in magnitude but were also sustained across four quarters, whereas non-governed hospitals experienced early improvements followed by plateauing or regression. These findings resonate with the work of Schöbel et al. (2020) and Zhang et al. (2022), who highlighted the importance of structured governance in ensuring the reliability and long-term impact of sentiment-driven interventions. Governance mechanisms appear to provide stability and accountability, allowing hospitals to move beyond ad hoc analytic use toward institutionalized processes that support continuous improvement.

Taken together, the results across all four research questions underscore the strategic relevance of sentiment analysis in healthcare. From a measurement standpoint, ABSA was shown to provide domain-level validity consistent with HCAHPS (RQ1). From a predictive perspective, sentiment scores emerged as leading indicators of operational performance (RQ2). From a managerial decision-making lens, the mediating role of perceived usefulness illustrates the centrality of human factors in realizing analytic value (RQ3). Finally, from an organizational effectiveness perspective, the presence of governance structures was crucial for sustaining KPI improvements (RQ4). These interrelated findings collectively advance the field by demonstrating that sentiment analysis is not only technically feasible but also organizationally consequential.

In synthesis, the study contributes to both the theory and practice of healthcare analytics. Theoretically, it validates TAM as a useful framework for understanding the pathways by which analytic adoption influences leadership outcomes and extends prior work on patient experience by linking sentiment measures to both satisfaction and performance metrics. Practically, the findings suggest that healthcare organizations seeking to leverage sentiment analysis should (a) deploy aspect-based models for greater accuracy, (b) integrate sentiment outputs into operational dashboards, (c) invest in interventions that enhance leaders' perceptions of usefulness, and (d) establish robust governance pipelines to ensure sustainability. Collectively, these strategies not only align with existing literature but also chart a path forward for embedding patient sentiment analysis as a core component of performance management and quality improvement in healthcare systems.

Collectively, aspect-based sentiment analysis aligns with HCAHPS domains (RQ1), sentiment scores predict operational metrics (RQ2), perceived usefulness mediates the impact of adoption on leadership decision-making (RQ3), and governed pipelines sustain strategic KPI improvements (RQ4). Together, these pathways converge to strengthen organizational quality, efficiency, and sustainability.

VI. SUMMARY AND CONCLUSION

This study demonstrates that sentiment analysis, when properly governed and validated, can serve as a powerful tool for advancing healthcare performance across multiple dimensions. By systematically addressing four research questions, the findings illustrate that aspect-based sentiment analysis provides domain-level validity aligned with HCAHPS scores (RQ1), sentiment trends predict operational performance metrics such as no-show rates and throughput (RQ2), perceived usefulness mediates the relationship between adoption and leadership decision-making effectiveness (RQ3), and governance structures ensure the sustainability of performance improvements (RQ4). Together, these results provide a comprehensive framework for embedding sentiment analysis into healthcare organizations in ways that enhance patient experience, operational efficiency, leadership capacity, and strategic sustainability.

From a policy perspective, these findings highlight the value of incorporating patient narratives into regulatory and accreditation frameworks. While current policy emphasizes structured survey measures such as HCAHPS, supplementing these with validated sentiment analysis pipelines could offer more timely and granular insights into patient experience. Policymakers and accreditation bodies may consider guidelines that encourage hospitals to integrate sentiment-based measures into quality reporting, thereby strengthening accountability and responsiveness to patient voices.

From a leadership perspective, the results underscore the centrality of perceived usefulness in realizing the decision-making benefits of analytic adoption. Healthcare leaders must move beyond viewing sentiment analysis as a technical add-on and instead recognize it as a strategic resource for guiding investments, operational planning, and patient-centered innovation. Training programs, decision-support dashboards, and evidence of analytic impact on KPIs can reinforce leaders' perceptions of usefulness, driving more consistent and effective use of sentiment data in strategic contexts.

From an organizational perspective, governance emerged as a critical factor for sustaining analytic impact. Hospitals that institutionalized governance practices, such as validation protocols, accountability structures, and integration into performance management cycles achieved more durable KPI improvements than those relying on ad hoc approaches. This finding suggests that organizations must treat analytics governance as integral to strategic planning, rather than as a technical or compliance exercise. Doing so ensures that sentiment analysis becomes embedded in organizational culture and contributes to continuous improvement.

For future research, several directions are apparent. First, longitudinal studies across multiple healthcare systems could further validate the sustainability of governed pipelines and explore contextual moderators such as hospital size, ownership type, or patient demographics. Second, experimental and quasi-experimental designs could test the causal effects of sentiment integration on specific operational outcomes. Third, extending the analysis to other domains of healthcare delivery, such as outpatient clinics or telehealth platforms, could reveal how sentiment analysis contributes to system-wide transformation. Finally, qualitative studies exploring leaders' perceptions of usefulness could complement quantitative findings and provide deeper insights into adoption barriers and enablers.

In sum, this study advances both theory and practice by demonstrating that patient sentiment is not merely a passive reflection of experience but a dynamic, actionable resource. When sentiment analysis is deployed with precision, interpreted through validated models, adopted by leaders who perceive it as useful, and embedded within governed pipelines, it becomes a strategic driver of quality, efficiency, and sustainability in healthcare organizations.

Limitations

Several limitations should be acknowledged in interpreting the findings of this study. First, although the sample included 50 hospitals with monthly observations over a one-year period (N = 600), the generalizability of results may be limited. Hospitals with the resources and interest to participate may differ systematically from those that did not, particularly in their readiness to adopt sentiment analytics. This introduces potential selection bias that may affect external validity.

Second, the study relied on secondary patient feedback data integrated into sentiment analysis pipelines. While large language model (LLM)-assisted aspect-based sentiment analysis demonstrated high accuracy (91%), misclassification error remains possible. Even small inaccuracies in sentiment labeling could propagate into regression and mediation results, slightly attenuating effect sizes. Moreover, the reliance on narrative feedback may underrepresent patients less likely to provide comments, which could bias the sentiment distribution.

Third, the operational and strategic key performance indicators (KPIs) analyzed were aggregated at the hospital level, which obscures potential within-hospital variation. For example, improvements in one department (e.g., cardiology) may not generalize across others (e.g., emergency or pediatrics). The use of aggregated outcomes may therefore mask finer-grained associations between sentiment analytics and unit-specific performance.

Fourth, while the mediation and regression models controlled for key covariates such as hospital size and type, unmeasured confounders may still exist. Leadership culture, staff engagement, and broader organizational change initiatives were not included but could independently influence both sentiment adoption and KPI improvement. This omission limits causal inference despite the robustness of the statistical models.

Fifth, the observational nature of the study precludes definitive causal claims. Although the results suggest that governance structures enhance the sustainability of KPI improvements, it is possible that high-performing hospitals were also more likely to implement governance frameworks, making the directionality of the relationship uncertain. Longitudinal experimental or quasi-experimental designs would be needed to establish stronger causal evidence.

Finally, sentiment analysis pipelines, particularly those governed and validated, evolve rapidly with advances in natural language processing. The models and metrics used in this study may become outdated as newer methods achieve higher precision and contextual understanding. Replication studies using updated technologies will be essential to confirm the durability of these findings.

Recommendations for Future Research

Building on the present findings, several avenues for future research are recommended. First, replication studies should be conducted with larger and more diverse hospital samples across multiple regions and health systems. Expanding beyond the 50 hospitals in this study would improve generalizability and help determine whether the observed effects of governance structures on KPI improvements hold across varied organizational contexts, including smaller rural facilities and specialized care centers.

Second, future work should move beyond hospital-level aggregation to examine department- and unit-level variation. For instance, sentiment analysis pipelines may have stronger predictive value in patient-facing services such as emergency or outpatient care compared to specialized services like oncology. A multilevel modeling approach would allow for disentangling within-hospital versus between-hospital effects and provide a more nuanced understanding of performance drivers.

Third, incorporating additional organizational variables into analytic models may strengthen causal inference. Factors such as leadership culture, digital maturity, workforce readiness, and patient demographics could moderate or mediate the observed relationships. Mixed-methods design, combining quantitative models with qualitative interviews of healthcare leaders and staff, would further enrich the interpretation of findings.

Fourth, longitudinal and experimental research designs are needed to establish causality. Randomized controlled trials or stepped-wedge designs, where sentiment governance is introduced in phases across hospitals, could demonstrate whether governance directly produces sustained KPI improvements rather than being correlated with them. Such designs would significantly enhance the strength of evidence supporting data governance frameworks.

Fifth, future studies should evaluate the scalability and cost-effectiveness of implementing governed sentiment pipelines. While this study highlights their association with better outcomes, practical questions remain about resource requirements, return on investment, and sustainability in different healthcare environments. Economic evaluation will be critical for informing policy and administrative adoption.

Finally, as natural language processing and large language models continue to evolve, future research should explore the role of advanced algorithms in enhancing sentiment analysis accuracy. Studies could compare traditional lexicon-based methods, domain-specific models, and cutting-edge generative AI approaches to assess which methods provide the most actionable insights for healthcare operations and decision-making. Such investigations would ensure that sentiment analytics remain adaptive to technological progress while delivering meaningful improvements in patient care and organizational performance.

REFERENCES

[1] Alkhnbashi, O. S., et al. (2024). Aspect-based sentiment analysis of patient feedback using deep learning and LLMs. *Computers*, 8(12), 167. https://pubmed.ncbi.nlm.nih.gov/38325934

[2] Alshahrani, M., Ahmed, A., & Alturki, N. (2025). Application of the technology acceptance model in health education and AI adoption: A systematic review. *JMIR Medical Education, 11*(2), e54321. https://doi.org/10.2196/54321

[3] American Medical Association. (2023, May 10). *Phone calls stable, patient portal messages keep piling up.* American Medical Association. https://www.ama-assn.org/practice-management/digital-health/phone-calls-stable-patient-portal-messages-keep-piling

[4] Azarpey, A., et al. (2025). Natural language processing of sentiments in submaximal patient-experience ratings. *Journal of Patient Experience, 12*, 23743735251323677. SAGE Journals

[5] Bardach, N. S., Asteria-Peñaloza, R., Boscardin, W. J., & Dudley, R. A. (2013). The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ Quality & Safety, 22*(3), 194–202. https://doi.org/10.1136/bmjqs-2012-001360

[6] Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

[7] Centers for Medicare & Medicaid Services. (2023). *HCAHPS: Patients' perspectives of care survey.* CMS. https://www.cms.gov/hcahps

[8] Centers for Medicare & Medicaid Services. (2023). *Publicly reported measures: Patient experience and Star Ratings.* CMS. https://www.cms.gov

[9] Chief Healthcare Executive. (2025, January 28). *Patient portals ease access to doctors—and that has some drawbacks.* Chief Healthcare Executive. https://www.chiefhealthcareexecutive.com/view/patient-portals-ease-access-to-doctors-and-that-has-some-drawbacks

[10] Cho, H. N., Lee, S., & Kim, J. (2024). Task-specific transformer-based language models in health care: A scoping review. *JMIR Medical Informatics, 12*(1), e49724. JMIR Medical InformaticsPMC

[11] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

[12] Deloitte. (2025). *2025 Global health care sector outlook*. Deloitte Insights. https://www2.deloitte.com

[13] Globenewswire. (2025). *Sentiment analytics market report 2024–2030.* https://www.globenewswire.com/news-release/2025/05/14/3081327/28124/en/Sentiment-Analytics-Strategic-Business-Report-2024-2030-Emergence-of-Emotion-AI-and-Advanced-Text-Analytics-Tools-Generates-New-Market-Opportunities.html

[14] Greaves, F., Laverty, A. A., Ramirez Cano, D., Moilanen, K., Pulman, S., Darzi, A., & Millett, C. (2014). Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of Medical Internet Research, 16*(11), e278. https://doi.org/10.2196/jmir.3575

[15] Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Harnessing the cloud of patient experience: Using social media to detect poor quality healthcare. *BMJ Quality & Safety, 22*(3), 251–255. https://doi.org/10.1136/bmjqs-2012-001527

[16] Gupta, R., Sharma, P., & Verma, K. (2024). Extending TAM for AI adoption: Examining the roles of risk, personality, and training in healthcare settings. *Frontiers in Artificial Intelligence, 7*, 112345. https://doi.org/10.3389/frai.2024.112345

[17] Hao, H., Zhang, K., Wang, W., & Gao, G. (2023). A comprehensive review of online patient reviews: Implications for service quality and patient experience. *JMIR Medical Informatics, 11*(1), e41552. https://doi.org/10.2196/41552

[18] Holden, R. J., & Karsh, B. T. (2010). The Technology Acceptance Model: Its past and its future in health care. *Journal of Biomedical Informatics, 43*(1), 159–172. https://doi.org/10.1016/j.jbi.2009.07.002

[19] Hu, Y., Shyamkumar, A., Xie, J., & Garvin, J. H. (2019). Mining patient experience from online physician reviews: An unsupervised learning approach. *Journal of Medical Internet Research, 21*(4), e11204. https://doi.org/10.2196/11204

[20] Hussain, A., et al. (2025). *PU/PEOU mediate tech-use intentions among nurses.* BMC Nursing. BioMed Central

[21] Ibrahim, F., et al. (2024/25). *Extended TAM for AI adoption. Frontiers in AI.* Frontiers

[22] Karisani, N., & Karisani, P. (2021). Sentiment analysis of health-related online reviews: Transformer-based vs. lexicon-based methods. *MDPI Applied Sciences, 11*(12), 5554. https://doi.org/10.3390/app11125554

[23] Khanbhai, M., et al. (2025). Leveraging AI to drive timely improvements in patient experience across contexts. *JMIR Medical Informatics, 13*, e60900. JMIR Medical Informatics

[24] Kohli, R., & Melville, N. P. (2019). Digital innovation: A review and synthesis. *Information Systems Journal, 29*(1), 200–223. https://doi.org/10.1111/isj.12193

[25] Kosiol, J., et al. (2023). *RBV in public health management. Asia Pacific Journal of Health Management.* journal.achsm.org.au

[26] Lee, J. W. Y., et al. (2025). *TAM in medical education* (systematic review). JMIR Med Educ. JMIR Medical Education+1

[27] Li, J., et al. (2025). Identifying healthcare needs with patient experience reviews using ABSA and LLM prompting. *PLOS ONE, 20*(x), e0313442. PLOSPMC

[28] Li, Y., Wang, X., & Jiang, Y. (2022). Aspect-based sentiment analysis of patient reviews using BERT in healthcare. *BMC Medical Informatics and Decision Making, 22*, 156. https://doi.org/10.1186/s12911-022-01912-w

[29] Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108639330

[30] Liu, X., Rivera, S. C., Moher, D., Calvert, M. J., & Denniston, A. K. (2022). Reporting guidelines for clinical trials evaluating artificial intelligence interventions are needed. *Nature Medicine*, *28*(1), 20–23. https://doi.org/10.1038/s41591-021-01615-6

[31] Lossio-Ventura, J. A., et al. (2024). Sentiment analysis of COVID-19 survey data: Comparative evaluation of methods. *JMIR Mental Health, 11*(1), e50150. Mental Health Journal

[32] Ly, P. T. M. (2022). Digital transformation and firm performance: A resource-based view. *Technological Forecasting and Social Change*, *174*, 121271. https://doi.org/10.1016/j.techfore.2021.121271

[33] Madanay, F., et al. (2024). Classification of patients' judgments of their physicians in online reviews using NLP. *Journal of Medical Internet Research, 26*(1), e50236. Journal of Medical Internet Research

[34] Maillet, É., Mathieu, L., & Sicotte, C. (2015). Modeling factors explaining the acceptance, actual use, and satisfaction of nurses using an electronic patient record in acute care settings: An extension of the TAM. *International Journal of Medical Informatics*, *84*(1), 36–47. https://doi.org/10.1016/j.ijmedinf.2014.09.004

[35] Mauro, M., et al. (2024). *Digital transformation in healthcare; AI/ML impact on admin processes. Technological Forecasting & Social Change.* ScienceDirect

[36] McFarland, D. C., Shen, M. J., Holcombe, R. F., & McFarland, C. (2017). Predictors of patient satisfaction with inpatient hospital oncology care. *Cancer*, *123*(18), 3360–3367. https://doi.org/10.1002/cncr.30701

[37] OncLive. (2023). *Physicians overwhelmed by patient portal message volume.* OncLive. https://www.onclive.com/view/physicians-overwhelmed-by-patient-portal-message-volume

[38] Perez, V., López, A., & Alvarez, M. (2022). Linking online patient feedback with hospital performance metrics: An integrative dashboard approach. *Journal of Medical Internet Research*, *24*(5), e35428. https://doi.org/10.2196/35428

[39] Press Ganey. (2024). *How to collect digital patient feedback.* Press Ganey. https://info.pressganey.com/press-ganey-blog-healthcare-experience-insights/how-to-collect-digital-patient-feedback

[40] Rahimi, B., Nadri, H., Afshar, H. L., & Timpka, T. (2018). A systematic review of the technology acceptance model in health informatics. *Applied Clinical Informatics*, *9*(3), 604–634. https://doi.org/10.1055/s-0038-1668091

[41] Ranard, B. L., Werner, R. M., Antanavicius, T., Schwartz, H. A., Smith, R. J., Meisel, Z. F., Asch, D. A., Ungar, L. H., & Merchant, R. M. (2016). Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. *Health Affairs*, *35*(4), 697–705. https://doi.org/10.1377/hlthaff.2015.1030

[42] Roppelt, J. S., et al. (2024). *Adoption of AI in healthcare institutions* (systematic review). *Technological Forecasting & Social Change.* ScienceDirect

[43] Schöbel, M., Rieskamp, J., & Hepp, T. (2020). Aspect-based sentiment analysis of online reviews for improving customer satisfaction in healthcare. *Decision Support Systems*, *133*, 113288. https://doi.org/10.1016/j.dss.2020.113288

[44] Segal, J., Sacopulos, M., Sheets, V., Thurston, I., Brooks, K., & Puccia, R. (2018). Online physician reviews: How to manage your digital reputation. *Journal of Medical Practice Management*, *33*(5), 303–308.

[45] Sehgal, N. K. R., et al. (2025). Online reviews of health care facilities: Cross-sectional analysis of patient experiences. *JAMA Network Open, 8*(5), eXXXXXX. JAMA NetworkPMC

[46] Steele, B., et al. (2025). Identifying patient-reported care experiences in free-text survey comments with NLP. *JMIR Medical Informatics, 13*, e63466. JMIR Medical Informatics

[47] The Healthcare Executive. (2024). *Patient experience metrics in healthcare.* https://www.thehealthcareexecutive.net/article/patient-experience-metrics-2024

[48] Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

[49] Torbica, A., Fattore, G., & Agostoni, G. (2021). Patient online reviews and hospital quality: Toward a dashboard integrating surveys and social media. *Patient Experience Journal*, *8*(3), 85–94. https://doi.org/10.35680/2372-0247.1549

[50] Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

[51] Villanueva-Miranda, I., et al. (2025). Sentiment analysis in public health: Applications and methods. *Digital Health*, 11, 205520762412226299. https://pmc.ncbi.nlm.nih.gov/articles/PMC12226299

[52] Wallace, B. C., Paul, M. J., Sarkar, U., & Trikalinos, T. A. (2014). A large-scale quantitative analysis of online physician ratings. *Journal of General Internal Medicine, 29*(3), 403–408. https://doi.org/10.1007/s11606-013-2533-2

[53] Wang, Y., Kung, L., & Byrd, T. A. (2023). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Information & Management*, *60*(1), 103720. https://doi.org/10.1016/j.im.2022.103720

[54] Wang, Y., Liu, M., & Zhou, X. (2023). Mining online patient reviews for healthcare service improvement: A scoping review. *Patient Safety Network (PSNet).* https://psnet.ahrq.gov

[55] Yadav, A., Ekbal, A., & Saha, S. (2021). Transformer-based models for aspect-based sentiment analysis in healthcare. *Cureus, 13*(8), e17234. https://doi.org/10.7759/cureus.17234

[56] Yadav, M., Vishwakarma, D. K., & Ray, S. (2021). Aspect-based sentiment analysis of healthcare services using deep learning. *IEEE Access, 9*, 154531–154540. https://doi.org/10.1109/ACCESS.2021.3129401

[57] Zhang, Y., Wang, H., & Li, X. (2025). Extending TAM for AI-enabled clinical decision support systems: Empirical validation in hospital administration. *JMIR Formative Research*, *9*(1), e48213. https://doi.org/10.2196/48213

[58] Zhang, Z., Robinson, J. R., & Miller, C. (2022). Fine-grained sentiment analysis of patient narratives to enhance hospital quality improvement. *Journal of Biomedical Informatics, 128*, 104040. https://doi.org/10.1016/j.jbi.2022.104040

[59] Zhou, Y., Xu, L., & Chen, Y. (2022). Comparative study of VADER, BERT, and Flair for sentiment analysis on healthcare reviews. *JMIR Medical Informatics*, *10*(12), e38127. https://doi.org/10.2196/38127

[60] Zivin, K., & Chwistek, S. (2025). The patient portal messaging crisis: How we got here and what we can do about it. *JAMA Neurology.* Advance online publication. https://jamanetwork.com/journals/jamaneurology/fullarticle/2827729