ORIGINAL RESEARCH **Open access**

Multimodal Neural Network for Predicting Patient Complaint Escalation Using Call Center Transcripts, Portal Messages, Service Recovery Notes, Encounter Metadata, and Prior Satisfaction Survey Responses

Jose Martinez[1*], Carmen Lopez[1]

Abstract

Escalated patient complaints signal breakdowns in communication, care coordination, service recovery, and trust. These events can create safety, reputational, financial, and legal risk, yet they often emerge from data already generated before escalation occurs. Current complaint management workflows are largely reactive. Health systems often rely on manual triage, retrospective reports, and delayed recognition of risk after a patient's dissatisfaction has already intensified. This article proposes a multimodal neural network for predicting whether an initial patient complaint is likely to escalate into a formal grievance, regulatory concern, legal threat, or other high-risk pathway. The model is conceptual and intended to support early intervention rather than replace human judgment. The proposed architecture integrates call center transcript embeddings, portal message language features, service recovery note representations, structured encounter metadata, and prior satisfaction survey responses. These inputs are fused into a dynamic escalation risk representation that could be updated as new interactions occur. Conceptually, the model would identify high-risk patterns such as repeated calls with increasing anger, portal messages expressing urgency, unresolved service recovery documentation, and prior low satisfaction scores. Such early detection could help patient relations teams prioritize timely outreach and targeted resolution. A multimodal neural network for complaint escalation prediction could shift patient relations from retrospective damage control toward proactive experience improvement. Its value would depend on careful validation, interpretability, privacy protection, and responsible integration into service recovery workflows.

Keywords Natural language processing, Multimodal deep learning, Patient complaints, Patient experience, Complaint escalation, Service recovery

*Correspondence:
Jose Martinez
jose.martinez@gmail.com
[1] Department of Health Informatics and Digital Analytics, Faculty of Medicine, University of Salamanca, Salamanca, Spain

Introduction

Escalated patient complaints represent more than dissatisfaction with hospitality or convenience; they can signal breakdowns in communication, clinical coordination, professionalism, safety, and organizational responsiveness. Prior research has shown that unsolicited patient complaints and observations may identify clinicians or care processes associated with adverse outcomes, malpractice exposure, or other forms of institutional risk [1-3]. Complaint escalation

therefore has implications for reputation, liability, quality improvement, and patient trust, particularly when early expressions of concern are not recognized as warning signals [4]. A predictive model that treats complaints as longitudinal risk events rather than isolated service episodes could help health systems intervene before dissatisfaction hardens into formal grievance or legal action [5].

Most patient relations workflows remain reactive, with escalation recognized only after a patient has made repeated contacts, requested leadership involvement, filed a formal grievance, or signaled legal intent. Machine learning work on complaint triage demonstrates that algorithmic tools can assist with classifying and prioritizing patient complaints, but much of this work still focuses on categorization after the complaint has already entered the system [5-7]. Reviews of complaint and misconduct risk prediction further suggest that health systems need earlier identification strategies that connect communication history, practitioner risk, and organizational response patterns [4]. The time lag between an initial concern and a formal escalation creates a critical window in which predictive analytics could support more responsive service recovery [8].

Modern healthcare organizations generate rich multimodal data before complaint escalation occurs. Call center transcripts can preserve the patient's words, issue framing, sentiment, and conversational trajectory, while portal messages and secure texts provide written signals of urgency, frustration, unresolved expectations, and repeated requests [9-12]. Service recovery notes capture how the organization responded, including whether explanations, apologies, compensatory actions, or follow-up commitments were documented, and structured encounter metadata can contextualize the complaint within visit type, unit, provider, length of stay, or discharge timing [13-15]. Prior satisfaction survey responses add another longitudinal signal, because low perceived quality, negative comments, and deteriorating experience scores may precede later complaints [16, 17].

This article proposes a conceptual multimodal neural network for predicting patient complaint escalation using call center transcripts, portal messages, service recovery notes, encounter metadata, and prior satisfaction survey responses. The model builds on advances in patient feedback NLP, patient portal message classification, complaint categorization, and multimodal healthcare data fusion. Rather than reporting experimental results, the article specifies an architecture that could integrate heterogeneous signals into a dynamic risk score and support proactive intervention by patient relations or risk management teams. The central thesis is that complaint escalation prediction should be modeled as a multimodal, temporal, and interpretable risk prediction problem rather than as a static text classification task.

Background

The lifecycle of patient complaints

Patient complaints often unfold through a lifecycle that begins with an informal expression of dissatisfaction and may progress through repeated calls, portal messages, leadership requests, formal grievances, regulatory filings, or legal claims. Factors that may drive escalation include unresolved clinical concerns, perceived disrespect, communication failures, unclear explanations, delayed follow-up, repeated transfers, and a mismatch between patient expectations and organizational response [1, 3, 18]. Conversely, escalation may be defused when service recovery is timely, specific, empathic, and visibly connected to corrective action. A lifecycle view is consistent with prior work showing that unsolicited patient observations and formal complaints can function as early signals of risk, rather than merely retrospective accounts of dissatisfaction [2, 4].

NLP of call center transcripts and portal messages

Natural language processing can transform unstructured patient communications into structured signals such as sentiment, emotion, topic, urgency, issue type, and unresolved concern. Patient portal message research has shown that secure messages can be classified by intent, clinical concern, and communication function, while deep learning approaches can represent message semantics beyond simple keyword rules [9-12]. Patient feedback NLP studies

similarly demonstrate that free-text comments and experience narratives can be mined for concerns, sentiment, dissatisfaction, and service quality signals [8, 19-22]. In the complaint escalation setting, call transcripts and portal messages would be expected to provide complementary linguistic evidence of whether a patient's concern is becoming more intense, repetitive, or adversarial [23, 24].

Service recovery notes as a window into resolution

Service recovery notes are internal records of how patient relations, managers, or clinical leaders responded to dissatisfaction, and they may contain signals that are not visible in the original complaint text. These notes can document whether an apology was offered, whether a clinical explanation was provided, whether a callback was promised, whether compensation or corrective action was discussed, and whether the patient accepted the response. Although complaint classification studies often focus on the patient's initial narrative, escalation risk also depends on whether the organization's response resolves or amplifies the concern [1, 6, 7]. NLP methods developed for patient feedback and patient concerns could be adapted to service recovery documentation to identify unresolved issues, vague closure language, missing follow-up, or repeated reopening of a case [19, 20, 22].

Encounter metadata and prior satisfaction as context

Encounter metadata provides context for interpreting complaint signals because escalation risk may vary by visit type, care setting, discharge timing, unit, procedure, clinician role, and the complexity of the patient's healthcare journey. Prior work linking patient observations, professional behavior, and adverse events suggests that complaint data can reflect deeper organizational or clinician-level risk patterns, especially when connected to structured clinical and operational information [1, 2, 18]. Satisfaction survey responses and patient comments can add further context by capturing perceptions of communication, responsiveness, coordination, respect, and overall care before the complaint episode [14-17]. A patient with prior low satisfaction, recent discharge, and multiple unresolved contacts may therefore carry a different escalation profile than a patient with a single low-intensity inquiry [13].

Multimodal learning for healthcare risk prediction

Multimodal learning is well suited to complaint escalation prediction because the target outcome depends on language, interaction history, structured encounter context, and organizational response. Reviews of biomedical and healthcare multimodal learning describe architectures that combine text, tabular variables, imaging, time series, audio, or other data streams into joint representations for prediction and decision support [25-27]. In patient experience, this principle suggests that portal messages, call transcripts, recovery notes, metadata, and survey scores should not be modeled as isolated silos. A fusion architecture could learn how weak signals across modalities combine into a stronger escalation trajectory, while still allowing each modality to contribute modality-specific evidence [8, 10, 25].

Model Development Overview

High-level predictive pipeline

The proposed pipeline would ingest each new patient interaction as it becomes available and update a dynamic complaint escalation risk score over time. A call center transcript, portal message, service recovery note, encounter update, or newly available satisfaction response would be processed by modality-specific encoders and passed to a temporal risk layer. If the estimated risk crossed a locally defined action threshold, the system could alert patient relations staff with a concise explanation and case summary, while preserving human review and discretion [5-7]. This design extends complaint triage from a static classification task into a streaming early-warning process that aligns with evidence that patient complaints and observations can identify broader quality and risk concerns [1, 4].

Core input modalities

The core input modalities include call transcript features derived from automatic speech recognition text and, where available, acoustic-prosodic cues such as speaking rate, pitch variation, pauses, interruptions, and rising agitation. Written modalities include portal messages and secure texts, represented with clinical language embeddings, as well as service recovery notes represented separately to distinguish patient expression from organizational response [9-12]. Structured modalities include encounter metadata such as admission date, discharge date, unit, provider role, visit type, length of stay, and time since discharge, while prior satisfaction survey data contribute global ratings, domain scores, and free-text experience comments [13-17]. Together, these inputs allow the model to represent both the patient's expressed concern and the institutional context in which escalation may unfold [8, 19].

Design principles

The model should be real-time updatable, privacy-compliant, interpretable to patient relations staff, and compatible with historical complaint data in which many cases resolve without formal escalation. Because complaint outcomes are not always observed immediately, the training framework would need to handle censored or ongoing cases in which the final escalation status is unknown at the time of model updating. The architecture should also avoid reducing complaint handling to automated denial or prioritization without context; instead, it should support earlier, more empathic, and better targeted human outreach [20-22]. These principles align with broader calls for explainable, validated, and workflow-integrated predictive systems in healthcare multimodal learning [25-27].

Data Sources and Feature Engineering

Call center transcripts and acoustic features

Call center data would be transformed into both text and interaction features, including patient language, staff language, speaker turns, interruptions, repeated issues, emotional intensity, and references to prior unresolved contacts. Automatic speech recognition outputs could be embedded with transformer-based language models, while acoustic representations could encode pitch, intensity, silence, hesitations, and speaking rate as proxies for distress or escalation tone. Voice-based clinical encounter research and call center sentiment analysis suggest that speech and transcript features can be used to assess quality, emotion, and interaction patterns, although complaint escalation applications would require careful validation in healthcare-specific settings [23, 24]. These features would be especially useful when written portal messages are absent but dissatisfaction is repeatedly expressed by telephone [8, 19].

Portal messages and service recovery notes

Portal messages would be processed for sentiment, urgency, dissatisfaction, request type, unresolved issue markers, repetition, legal language, and references to safety, billing, disrespect, or delayed communication. Existing patient portal message classification studies support the feasibility of using machine learning and deep learning to identify message intent, content categories, and patient concerns from secure communications [9-12]. Service recovery notes would be encoded separately to represent the organization's response, including apology, explanation, escalation to leadership, commitment to follow-up, documentation of closure, and evidence that the patient remained dissatisfied. Combining patient-authored and staff-authored text would allow the model to distinguish an emotionally intense complaint that receives effective recovery from a moderate complaint that remains unresolved and therefore becomes more likely to escalate [1, 6, 20].

Encounter metadata and satisfaction survey integration

Structured encounter metadata would be converted into a fixed-length feature vector containing variables such as visit setting, service line, admission and discharge timing, unit, provider role, encounter duration, and time elapsed since the relevant care episode. Prior satisfaction survey responses would contribute numerical domain scores, overall ratings, and patient comment embeddings, capturing the patient's broader experience trajectory before the complaint event [14-17]. Machine learning studies of patient satisfaction and perceived service quality suggest that structured ratings and free-text

feedback can reveal dissatisfaction patterns that may not appear in formal complaint databases [13, 16, 17]. Integrating metadata and surveys with communication data would help the model contextualize whether a complaint reflects an isolated issue, a cumulative negative experience, or a higher-risk pattern associated with repeated dissatisfaction [1, 4].

Table 1 consolidates the proposed model's multimodal input structure by linking each data source to its representation strategy, escalation-risk contribution, and practical interpretation for patient relations staff.

Table 1. Multimodal Input-to-Representation Framework for Patient Complaint Escalation Prediction

Input modality	Primary operational signal	Example raw data elements	Representation strategy	Escalation-risk contribution	Practical interpretation for patient relations
Call center transcripts	Spoken dissatisfaction, repeated contact, tone of interaction, unresolved issue framing	Transcript text, speaker turns, interruptions, repeated calls, references to prior unresolved contacts	Transformer-based transcript embeddings combined with interaction-sequence features	Detects escalating language, repeated attempts to resolve the same issue, increasing frustration, and adversarial wording	Identifies cases where telephone communication suggests growing dissatisfaction even before formal grievance language appears
Acoustic and prosodic call features	Emotional intensity and interaction stress not fully captured in transcript text	Speaking rate, pauses, pitch variation, volume, interruptions, silence, agitation markers	Sequence or convolutional acoustic encoder linked to call-level features	Adds non-textual evidence of distress, urgency, frustration, or breakdown in communication	Helps staff recognize emotionally intense calls that may require rapid empathic outreach
Portal messages and secure texts	Written urgency, dissatisfaction, repetition, legal or safety concern language	Portal message content, repeated messages, urgent phrases, unresolved requests, safety/billing/disrespect references	Clinical language model or transformer-based message embeddings	Captures persistent written concern, urgency escalation, and explicit dissatisfaction across asynchronous communication	Supports prioritization of patients whose written messages indicate unresolved or intensifying concern
Service recovery notes	Organizational response quality and documented resolution status	Apology, explanation, callback promise, leadership review, compensation, closure note, unresolved case language	Separate staff-authored note encoder to distinguish recovery documentation from patient-authored communication	Indicates whether the health system response may reduce or fail to reduce escalation risk	Helps differentiate a severe complaint with strong recovery from a moderate complaint with weak or vague closure

Encounter metadata	Clinical and operational context surrounding the complaint	Unit, service line, visit type, discharge timing, provider role, length of stay, time since encounter	Fixed-length structured feature vector processed through feed-forward layers	Contextualizes whether the complaint follows a complex, high-risk, delayed, or sensitive care episode	Gives staff relevant operational context before outreach or case review
Prior satisfaction survey responses	Longitudinal patient experience trajectory before the complaint	Overall rating, domain scores, communication ratings, free-text comments, prior negative experience	Numerical score encoding plus text-comment embeddings	Identifies cumulative dissatisfaction and prior low trust that may increase escalation risk	Alerts staff when the current complaint may be part of a broader negative experience pattern
Cross-modal temporal sequence	Escalation trajectory over time rather than isolated complaint content	Chronological calls, messages, recovery notes, encounter updates, new survey responses	Time-aware aggregation layer, recurrent model, temporal transformer, or survival-oriented extension	Updates risk as new events occur and distinguishes improving, unresolved, or worsening trajectories	Supports dynamic case monitoring instead of one-time complaint classification

Multimodal Neural Network Architecture

Modality-specific encoders

The proposed architecture would use separate encoders for each modality so that the model can preserve the structure of different data streams before fusion. Call transcript text could be processed with a transformer encoder, acoustic features with a convolutional or sequence model, portal messages with a clinical language model, service recovery notes with a distinct text encoder, and tabular encounter and survey variables with a feed-forward network [9, 10, 24]. This separation is important because patient-authored messages, staff-authored notes, acoustic signals, and structured metadata carry different kinds of evidence about escalation risk. Multimodal healthcare reviews support modality-specific representation learning as a foundation for later fusion into a shared predictive space [25-27].

Fusion layer and joint representation

After modality-specific encoding, the model would combine representations through late fusion, gated fusion, or cross-modal attention to produce a joint latent representation of the complaint trajectory. Simple concatenation could provide a transparent baseline architecture, while attention-based fusion could allow the model to learn when one modality, such as a highly urgent portal message or a service recovery note documenting unresolved issues, should receive greater weight [25, 26]. This joint representation would be expected to capture patterns that no single modality can reveal alone, such as the combination of a poor prior survey, multiple angry calls, and a vague closure note [13-15]. Such fusion is consistent with patient feedback research showing that narrative, structured, and temporal information can each contribute to understanding dissatisfaction and service quality [8, 19, 22].

Figure 1 illustrates the proposed multimodal complaint-escalation workflow, showing how call transcripts, portal messages, service recovery notes, encounter metadata, and prior satisfaction responses are transformed into a temporally updated escalation-risk estimate for human-led service recovery.

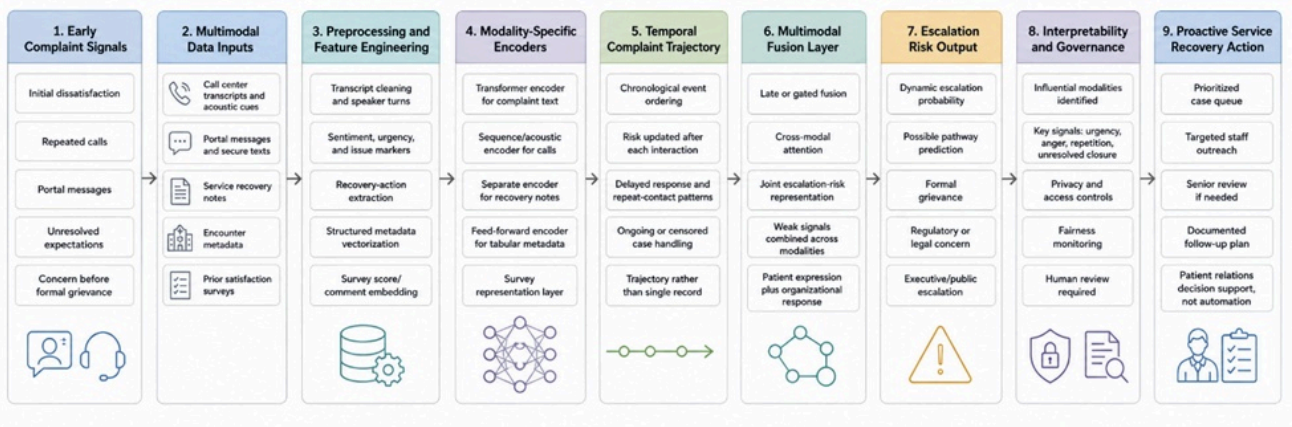


Figure 1. Multimodal Neural Network Workflow for Predicting and Prioritizing Patient Complaint Escalation in Service Recovery Operations

Escalation risk output

The final layer would output a continuously updated escalation risk probability, which patient relations teams could interpret as a prioritization aid rather than a definitive judgment. Additional conceptual heads could be designed to estimate likely escalation pathways, such as formal grievance, regulatory complaint, legal threat, executive complaint, or public social media escalation, but these outputs would require careful validation and governance before operational use [1, 6, 7]. The model should also produce an explanation layer that identifies influential modalities and signals, such as repeated urgent messages, lack of documented resolution, deteriorating sentiment, or prior low satisfaction [14, 20, 21]. In practice, the output should be calibrated for workflow actionability, supporting timely service recovery while avoiding automated decisions that could intensify patient mistrust [1, 4].

Incorporating Temporal Dynamics and Escalation Pathways

Temporal sequence of interactions

Complaint escalation is inherently temporal because dissatisfaction may accumulate across calls, portal messages, callbacks, documentation updates, and unresolved recovery attempts. The model would therefore process events in chronological order, using a recurrent, transformer-based, or time-aware aggregation layer to update the escalation risk score whenever a new interaction occurs [25-27]. This temporal structure would allow the network to represent patterns such as increasing emotional intensity across repeated calls, delayed response after discharge, or a portal message that reopens an issue previously marked as resolved [9, 10, 12]. Rather than treating a complaint as a single static text record, the architecture would model the complaint as a trajectory in which risk rises or falls as the patient and organization continue to interact [4, 5].

Handling censored and on-going cases

Many complaint episodes remain unresolved for a period of time, and some never escalate despite appearing high risk early in the process. A conceptually appropriate model would therefore account for censored cases, in-progress service recovery, and variable observation windows, rather than assuming that every case has a fully observed endpoint [4, 27]. Survival-oriented extensions or time-to-event framing could help represent whether escalation is likely within a future period while preserving the distinction between a resolved case and a case that has not yet had enough follow-up time [1, 3]. This is especially important in patient relations workflows, where a case may remain open while additional information, leadership review, or patient callback is pending [5, 6].

Feedback loop from service recovery actions

Service recovery actions should not be treated merely as documentation after the complaint, because they may actively change the probability of escalation. The model could condition its temporal risk estimate on the type, timing, and specificity of recovery actions, including apology, explanation, care team review, billing correction, leadership escalation, or documented follow-up [6, 7, 20]. If risk remains high or increases after a recovery note claims closure, the model could flag a possible mismatch between organizational documentation and the patient's continuing dissatisfaction [19, 22]. This feedback loop would support adaptive recovery, in which patient relations teams revise their intervention strategy when the available signals suggest that the initial response has not restored trust [8, 21].

Model Interpretability and Stakeholder Trust

Explaining escalation risk to patient relations staff

For patient relations staff, an escalation score would be useful only if accompanied by explanations that connect the prediction to recognizable case details. Modality-specific attribution methods, attention summaries, or post hoc explanation tools could identify leading contributors such as repeated contacts, urgent language, angry sentiment, absence of documented resolution, low prior satisfaction, or proximity to discharge [14, 20, 21]. These explanations should be written in operational language that supports empathetic outreach rather than in purely technical model terms [8, 19]. The goal would be to help staff understand why a case should be prioritized and what action might reduce escalation risk, while preserving human review of context and nuance [5, 7].

Transparency for risk management and legal

Risk management and legal stakeholders would need a transparent view of how escalation risk changes over time and which signals contribute to that trajectory. A dashboard could display the sequence of calls, messages, recovery notes, encounter events, and prior satisfaction information in relation to the evolving risk estimate, without presenting the model output as a legal conclusion [1, 2, 18]. This approach would support proactive resolution by highlighting preventable service failures, repeated communication breakdowns, or unresolved safety concerns before they become formal claims [3, 4]. Transparency would also make it easier to audit whether the model overweights certain types of language, service lines, or patient groups, which is essential for responsible deployment [25, 26].

Operational Deployment and Proactive Service Recovery

Real-time alerting and case prioritization

In deployment, the model would operate as a prioritization layer that updates risk when a new call transcript, portal message, recovery note, encounter update, or survey response enters the system. When risk exceeds a locally governed threshold, patient relations staff could receive an alert containing a concise summary of the case, key contributing signals, and recommended next steps for review [5-7]. The alert should not replace staff judgment, but it could help identify cases that might otherwise be missed because they are distributed across telephone logs, portal inboxes, survey systems, and service recovery documentation [9, 12, 13]. Such prioritization would be especially valuable when complaint volume is high and manual triage cannot consistently detect subtle escalation trajectories [8, 19].

Integration with crm and service recovery workflows

The model would be most useful if embedded directly within the hospital's patient relations, customer relationship management, or complaint management workflow. Integration could allow high-risk cases to be assigned to senior staff, prompt timely callbacks, document planned actions, and track whether subsequent patient communications indicate improvement or continuing dissatisfaction [6, 20, 22]. Structured encounter and survey data could enrich this workflow by

giving staff context about the care episode, prior experience ratings, and service line history before outreach occurs [14-17]. Operationally, the system should be designed as a service recovery support tool that helps organizations act earlier and more consistently, not as an automated gatekeeper for complaints [1, 4].

Evaluation Strategy

Predictive performance metrics

The model should be evaluated with discrimination, calibration, and timeliness metrics that are appropriate for escalation prediction, while avoiding reliance on a single summary measure. Conceptually, evaluation could include classification performance for escalation, precision-recall behavior for rare high-risk events, calibration of predicted risk, and assessment of whether risk becomes actionable before formal escalation occurs [5-7]. Temporal evaluation is important because a model that identifies risk only after repeated late-stage contacts would have limited operational value, even if it appears accurate retrospectively [4, 27]. Model assessment should also examine whether each modality contributes useful information beyond the others, consistent with multimodal learning principles [25, 26].

Temporal and external validation

Validation should be structured around future-facing use, because complaint practices, portal adoption, call center workflows, and service recovery documentation may shift over time. A forward-time validation design would test whether a model trained on earlier complaint episodes can support later case prioritization without learning artifacts from future documentation [4, 27]. External validation in another hospital, service line, or patient population would be needed before broad deployment, especially because complaint language, survey instruments, staffing models, and documentation norms differ across settings [8, 13, 19]. Validation should also compare multimodal fusion against simpler text-only or tabular-only approaches to determine whether the added complexity is justified [9, 10, 25].

Operational impact

Operational evaluation should examine whether the model changes service recovery behavior in ways that patients and staff experience as helpful. Relevant conceptual outcomes include earlier outreach, more consistent case prioritization, fewer unresolved repeat contacts, improved staff situational awareness, and stronger alignment between complaint handling and quality improvement [14, 20, 22]. Health systems may also assess whether proactive intervention is associated with fewer formal grievances, reduced downstream risk management burden, or better patient relations workflow efficiency, while recognizing that such outcomes require careful prospective study rather than assumed benefit [1, 3, 4]. Staff feedback should be incorporated because alert burden, explanation quality, and perceived usefulness will determine whether the model is trusted or ignored [7, 8].

Table 2 presents a governance and validation framework for determining whether the proposed multimodal escalation model is accurate, timely, interpretable, fair, privacy-preserving, and operationally ready for patient relations workflows.

Table 2. Governance, Validation, and Workflow Readiness Framework for Multimodal Complaint Escalation Prediction

Readiness domain	Key analytical question	Required evaluation or safeguard	Why it matters for this model	Operational decision-use implication
Predictive discrimination	Can the model distinguish complaints that escalate from those	Evaluate discrimination using appropriate classification and rare-event metrics, including	Escalated grievances may be relatively uncommon, making overall accuracy insufficient	The model should support prioritization of high-risk cases without overwhelming staff with low-value alerts

	that resolve without escalation?	sensitivity to high-risk cases		
Calibration	Do predicted risk probabilities correspond to observed escalation likelihood?	Assess calibration across risk strata, service lines, complaint channels, and time periods	Poorly calibrated risk scores could cause overreaction or underreaction by patient relations teams	Staff should be able to interpret risk levels as actionable priority signals, not arbitrary scores
Timeliness	Does the model identify risk early enough for service recovery to occur before formal escalation?	Measure lead time between model alert and later grievance, regulatory concern, legal threat, or executive complaint	A model that predicts escalation only after late-stage signals has limited operational value	Alerts should occur during the service recovery window, when outreach can still change the complaint trajectory
Incremental value of multimodal fusion	Does combining modalities improve usefulness beyond text-only or tabular-only models?	Compare multimodal architecture against call-only, portal-only, recovery-note-only, and metadata-only baselines	Multimodal complexity is justified only if it adds interpretable and practical value	Health systems should adopt fusion models only when additional data streams improve prioritization or explanations
Interpretability	Can staff understand why a case is being prioritized?	Provide modality-level and signal-level explanations, such as repeated contacts, urgent language, low satisfaction, or unresolved recovery notes	Patient relations staff need explanations that translate model output into empathic action	The system should generate case-review cues, not opaque risk labels
Fairness and representational equity	Does the model detect escalation risk consistently across patient groups and communication modes?	Audit predictions by language, demographic group, disability status, service line, portal use, and communication channel	Patients who complain through less-documented channels may be underrepresented	Deployment should avoid privileging digitally traceable dissatisfaction while missing less visible distress
Privacy and access control	Are sensitive complaint narratives protected from inappropriate secondary use?	Apply role-based access, audit trails, minimum necessary data use, de-identification where appropriate, and governance review	Complaint data may include protected health information, legal concerns, emotional narratives, and sensitive family context	The model should function as governed decision support within patient relations, not as unrestricted institutional surveillance
Workflow integration	Does the model fit existing patient relations and service recovery processes?	Pilot within CRM or complaint-management systems; assess alert burden, actionability, staff trust, and documentation quality	Even accurate models fail when alerts are poorly timed, poorly explained, or disconnected from staff workflow	Risk outputs should feed into case assignment, callback planning, escalation review, and documented follow-up

Prospective impact	Does model-supported service recovery improve outcomes without creating harm?	Conduct prospective evaluation of outreach timeliness, repeat contacts, grievance rates, staff workload, and patient experience	Conceptual benefit cannot be assumed without real-world testing	Implementation should proceed through monitored pilots before broad deployment

Limitations

Data availability and privacy

A key limitation is that the proposed model depends on data streams that may be incomplete, inconsistently documented, or governed by different privacy and consent requirements. Call recordings and automatic speech recognition transcripts may not be available for all patients, portal messages may omit dissatisfaction expressed by phone or in person, and service recovery notes may vary substantially in detail and tone [9, 11, 24]. These data sources may contain protected health information, sensitive personal narratives, legal concerns, and emotionally charged language, requiring strict access control, de-identification where appropriate, auditability, and governance [8, 19, 25]. The model should therefore be developed as a privacy-preserving decision support system with clear limits on secondary use and careful monitoring of how predictions influence staff behavior [26, 27].

Generalizability and fairness

Generalizability is uncertain because complaint expression differs across languages, cultures, disability status, communication channels, health literacy levels, and trust in the healthcare system. Patients who express dissatisfaction verbally during an encounter, through family members, or outside formal communication systems may be underrepresented, creating risk that the model privileges traceable digital dissatisfaction over less documented forms of distress [4, 8, 22]. Fairness evaluation should examine whether predictions, explanations, and alerting patterns differ across demographic groups, service lines, language preferences, and communication modalities [14, 17, 26]. Without such evaluation, a complaint escalation model could unintentionally amplify existing inequities in whose concerns are detected early and whose concerns remain invisible [1, 2].

Conclusion

A multimodal neural network for predicting patient complaint escalation would bring together the communication, documentation, encounter, and satisfaction signals that often precede formal grievances. By representing complaints as evolving trajectories rather than isolated records, the model could help patient relations teams identify cases that require earlier and more focused intervention.

The strongest feature of this approach is its ability to combine multiple weak signals into a more complete view of risk. Call transcripts, portal messages, service recovery notes, encounter metadata, and prior survey responses each provide only a partial view, but together they could support dynamic updating, interpretable explanations, and more timely service recovery.

Important challenges remain before such a model could be responsibly implemented. Health systems would need consistent data capture, strong privacy protections, careful governance, fairness monitoring, and prospective validation that demonstrates real operational value without increasing burden or eroding trust.

Pilot deployment in patient relations departments would be a practical next step. Such pilots should evaluate whether multimodal risk prediction helps staff intervene earlier, resolve concerns more effectively, and improve the patient

experience while reducing preventable escalation.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 26 Aug 2025 Revised: 23 Sep 2025 Accepted: 06 Nov 2025
Published online: 25 February 2026

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

Cooper WO, Guillamondegui O, Hines OJ, Hultman CS, Kelz RR, Shen P, et al. Use of unsolicited patient observations to identify surgeons with increased risk for postoperative complications. JAMA Surg. 2017;152(6):522-9.

Cooper WO, Martinez W, Domenico HJ, Callahan ST, Kirkby BP, Finlayson AJ, et al. Unsolicited patient complaints identify physicians with evidence of neurocognitive disorders. Am J Geriatr Psychiatry. 2018;26(9):927-36.

Grandizio LC, Barreto Rocha DF, Hayes D, Warnick EP, Doyle CM, Suk M, et al. An analysis of formal patient complaints, risk, and malpractice events involving orthopedic trauma surgeons during a 10-year period. Orthopedics. 2023;46(2):121-7.

Wang Y, Ram S, Scahill S. Risk identification and prediction of complaints and misconduct against health practitioners: a scoping review. Int J Qual Health Care. 2024;36(1):mzad114.

Elmessiry A, Cooper WO, Catron TF, Karrass J, Zhang Z, Singh MP. Triaging patient complaints: Monte Carlo cross-validation of six machine learning classifiers. JMIR Med Inform. 2017;5(3):e19.

Li X, Shu Q, Kong C, Wang J, Li G, Fang X, et al. An intelligent system for classifying patient complaints using machine learning and natural language processing: development and validation study. J Med Internet Res. 2025;27:e55721.

Koh SW, Wong ER, Tan JC, van der Lubbe SC, Goh JC, Ching ES, et al. Classifying patient complaints using artificial intelligence–powered large language models: cross-sectional study. J Med Internet Res. 2025;27:e74231.

Khanbhai M, Anyadi P, Symons J, Flott K, Darzi A, Mayer E. Applying natural language processing and machine learning techniques to patient experience feedback: a systematic review. BMJ Health Care Inform. 2021;28(1):e100262.

Cronin RM, Fabbri D, Denny JC, Rosenbloom ST, Jackson GP. A comparison of rule-based and machine learning approaches for classifying patient portal messages. Int J Med Inform. 2017;105:110-20.

Sulieman L, Gilmore D, French C, Cronin RM, Jackson GP, Russell M, et al. Classifying patient portal messages using convolutional neural networks. J Biomed Inform. 2017;74:59-70.

Tafti AP. Probing patient messages enhanced by natural language processing: a top-down message corpus analysis. Health Data Sci. 2021;2021:9876543.

Ren Y, Wu Y, Fan JW, Khurana A, Fu S, Wu D, et al. Automatic uncovering of patient primary concerns in portal messages using a fusion framework of pretrained language models. J Am Med Inform Assoc. 2024;31(8):1714-24.

Bari V, Hirsch JS, Narvaez J, Sardinia R, Bock KR, Oppenheim MI, et al. An approach to predicting patient experience through machine learning and social network analysis. J Am Med Inform Assoc. 2020;27(12):1834-43.

Liu N, Kumara S, Reich E. Gaining insights into patient satisfaction through interpretable machine learning. IEEE J Biomed Health Inform. 2020;25(6):2215-26.

Matsuda S, Ohtomo T, Okuyama M, Miyake H, Aoki K. Estimating patient satisfaction through a language processing model: model development and evaluation. JMIR Form Res. 2023;7:e48534.

Khaleghparast S, Maleki M, Hajianfar G, Soumari E, Oveisi M, Golandouz HM, et al. Development of a patients' satisfaction analysis system using machine learning and lexicon-based methods. BMC Health Serv Res. 2023;23(1):280.

Rahim AI, Ibrahim MI, Musa KI, Chua SL, Yaacob NM, et al. Assessing patient-perceived hospital service quality and sentiment in Malaysian public hospitals using machine learning and Facebook reviews. Int J Environ Res Public Health. 2021;18(18):9912.

Cooper WO, Spain DA, Guillamondegui O, Kelz RR, Domenico HJ, Hopkins J, et al. Association of coworker reports about unprofessional behavior by surgeons with surgical complications in their patients. JAMA Surg. 2019;154(9):828-34.

Feizollah A, Lin CY, O'Malley L, Thompson W, Listl S, Byrne M, et al. The use of natural language processing to interpret unstructured patient feedback on health services: scoping review. J Med Internet Res. 2025;27:e72853.

Shankar R, Yip A. Transforming patient feedback into actionable insights through natural language processing: knowledge discovery and action research study. JMIR Form Res. 2025;9:e69699.

Azarpey A, Thomas J, Ring D, Franko O. Natural language processing of sentiments identified in patient comments associated with less than top-rated care. J Patient Exp. 2025;12:23743735251323677.

Fairie P, Zhang Z, D'Souza AG, Walsh T, Quan H, Santana MJ, et al. Categorising patient concerns using natural language processing techniques. BMJ Health Care Inform. 2021;28(1):e100274.

Aattouri I, Mouncif H, Rida M. Call center customer sentiment analysis using ML and NLP. In: 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA). IEEE; 2023.

Habib M, Faris M, Qaddoura R, Alomari M, Alomari A, Faris H, et al. Toward an automatic quality assessment of voice-based telemedicine consultations: a deep learning approach. Sensors. 2021;21(9):3279.

Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review. Brief Bioinform. 2022;23(2):bbab569.

Krones F, Marikkar U, Parsons G, Szmul A, Mahdi A. Review of multimodal machine learning approaches in healthcare. Inf Fusion. 2025;114:102690.

Teles AS, de Moura IR, Silva F, Roberts A, Stahl D. EHR-based prediction modelling meets multimodal deep learning: a systematic review of structured and textual data fusion methods. Inf Fusion. 2025;118:102981.