

ORIGINAL RESEARCH

Open access

Machine Learning Model for Predicting High Administrative Burden Encounters Using Documentation Complexity, Billing Requirements, Insurance Rules, Care Coordination Needs, and Provider Workload Indicators

Mohamed Salah^{1*}, Youssef Karim¹, Ahmed Nabil², Mahmoud Adel¹, Karim Hassan²

Abstract

Administrative tasks surrounding a clinical encounter include documentation, coding, billing, insurance verification, prior authorization, and care coordination. These tasks are unevenly distributed across encounters and can consume substantial clinical and operational capacity. Health systems often detect administrative overload only after coding backlogs, payer denials, unanswered messages, or staff overtime have already emerged. The absence of an encounter-level prediction tool limits the ability of practices to intervene before administrative work accumulates. This article proposes a machine learning model that predicts whether an encounter is likely to become a high-administrative-burden event. The model uses documentation complexity, billing requirements, insurance rules, care coordination needs, and provider workload indicators as core predictors. A gradient-boosted classification framework is conceptually specified using historical encounter, billing, scheduling, payer, and workload data. The model would generate an encounter-level burden risk score and provide interpretable feature-domain contributions to support operational decisions. Conceptually, the model could identify encounters likely to require additional coding review, prior authorization follow-up, payer documentation, or multidisciplinary coordination. The resulting risk score would support proactive staffing, pre-visit review, and workflow routing. A predictive model for high administrative burden encounters could help shift healthcare administration from reactive queue management to anticipatory operational planning. Such a model may support revenue integrity, reduce avoidable rework, and lessen administrative strain on clinicians and staff.

Keywords Machine learning, Predictive analytics, Administrative burden, Revenue cycle, Documentation complexity, Provider workload

*Correspondence:

Mohamed Salah
mohamed.salah@gmail.com

¹ Department of Intelligent Healthcare Systems, Faculty of Medicine, Cairo University, Cairo, Egypt

² Department of Clinical Data Engineering, Faculty of Engineering, Ain Shams University, Cairo, Egypt

Introduction

Administrative burden has become a defining operational challenge in healthcare delivery because it links clinical documentation, revenue cycle processing, payer compliance, and clinician work conditions into a single cumulative load. Electronic health record activity studies show that physicians spend substantial time on documentation and desktop medicine, often outside the visible encounter itself, which connects encounter complexity to downstream professional

strain [1, 2]. Waste analyses and policy discussions further identify administrative complexity as a major contributor to health system inefficiency, with billing, insurance navigation, and compliance-related work consuming resources that could otherwise support patient care [3-5]. A predictive model that anticipates high-burden encounters would therefore address not only coding and billing delays but also the broader operational conditions associated with burnout and workflow fragmentation [6, 7].

Administrative burden varies considerably across encounters because the same visit type can produce different levels of documentation, coding specificity, payer review, and care coordination depending on patient, clinician, payer, and practice context. Encounters involving complex diagnoses, detailed notes, coding modifiers, plan-specific documentation expectations, or recent transitions of care are more likely to trigger administrative follow-up than routine visits with straightforward billing paths [8-10]. Provider workload also modifies encounter burden, because a documentation-intensive visit occurring during a dense clinic session or high in-basket period may impose greater operational strain than the same visit under less congested conditions [11-13]. This variability suggests that administrative burden should be modeled at the encounter level rather than assumed from specialty, visit type, or payer category alone.

Healthcare organizations increasingly possess the electronic data needed to anticipate administrative workload before it becomes visible as a backlog. EHR logs, note metadata, problem lists, diagnosis counts, coding histories, scheduling systems, and billing records can be linked to represent both the clinical and administrative context of an encounter [1, 14, 15]. Automated clinical coding research shows that structured and unstructured clinical information can support prediction of coding requirements, while revenue cycle analytics suggests that claim denials and coding accuracy are amenable to algorithmic risk assessment [16-19]. These data sources create the foundation for a model that would combine documentation complexity, billing requirements, insurance rules, coordination needs, and provider workload into a single prospective burden signal.

The thesis of this manuscript is that a predictive model could forecast administrative burden risk at the encounter level and enable proactive resource allocation before work accumulates downstream. Rather than treating denials, coding queries, documentation rework, and staff overtime as isolated events, the proposed model conceptualizes them as observable outcomes of a shared burden-generating process. A model that operates at scheduling, check-in, or early post-visit stages could support targeted coding review, payer verification, documentation prompts, and care coordination routines. In this way, predictive analytics could become an operational bridge between clinical documentation improvement, revenue integrity, and workload management.

Background

Administrative burden in healthcare delivery

Administrative burden in healthcare delivery refers to the cumulative work required to document care, justify services, code encounters, satisfy payer rules, coordinate follow-up, and resolve billing or authorization issues. Studies of EHR use demonstrate that administrative and documentation work is tightly embedded in clinical practice, while broader analyses of waste and administrative expense show that such work contributes materially to health system inefficiency [1-5]. Documentation burden has also been conceptualized as a measurable phenomenon, including effort related to note creation, review, inbox management, coding support, and compliance tasks [20]. For predictive modeling, administrative burden should therefore be treated as a multidomain construct rather than a single billing event.

Documentation complexity and billing requirements

Documentation complexity can be inferred from note length, unstructured clinical content, number of problems addressed, diagnostic specificity, and the degree to which a record supports coding and payer review. Automated clinical coding studies demonstrate that clinical text, ICD mappings, and structured diagnosis information can be used to infer coding needs, although interpretability and local workflow fit remain important considerations [9, 10, 16, 17]. Billing requirements

add a second layer of complexity because CPT or HCPCS coding level, modifier use, bundled services, and payer-specific edits may alter how much documentation and review are required for an otherwise similar encounter [18, 19]. A predictive burden model would need to represent both the clinical complexity of the note and the administrative complexity of converting that note into a compliant claim.

Insurance rules and care coordination needs

Insurance rules create administrative burden when prior authorization, medical necessity review, eligibility verification, plan-specific documentation, or appeal preparation is required. Predictive approaches to claim denials and health insurance decision support indicate that payer-facing administrative events can be modeled from available clinical and billing information, although such models must be adapted to local payer rules and documentation practices [18, 21]. Care coordination needs further increase burden when encounters involve multiple specialists, recent hospitalization, referrals, care gaps, or multidisciplinary follow-up that requires communication across organizational boundaries. In the proposed model, insurance rule indicators and coordination features would be treated as interacting drivers because payer requirements often become more difficult to satisfy when care pathways are clinically complex.

Provider workload indicators

Provider workload indicators capture the operational context in which the encounter occurs, including daily schedule density, patient volume, time of day, day of week, in-basket burden, documentation carryover, and available team support. EHR log studies show that physician workload is not limited to face-to-face visit time and that documentation and inbox activity can extend substantially beyond scheduled care sessions [1, 2, 14]. Research linking workload, burnout, errors, and EHR usability suggests that administrative burden should be modeled as a function of both encounter complexity and clinician-system capacity [6, 7, 11, 22]. A high-complexity encounter may be manageable in a well-supported schedule but become high burden when it coincides with accumulated documentation, dense clinic flow, or reduced staffing.

Machine learning for operational and revenue cycle analytics

Machine learning has been applied to clinical coding, claim denial prediction, hospital admission denial triage, and health system-scale prediction, showing that operational and administrative signals can be extracted from EHR and billing data [18, 21, 23-25]. Automated coding and ICD prediction studies provide methodological foundations for representing clinical text, structured diagnosis data, and coding requirements in predictive models [9, 15, 16, 19]. However, most existing approaches focus on narrower targets such as code assignment, denial likelihood, or documentation support rather than a holistic encounter-level burden construct. The proposed model addresses this gap by treating burden as a composite operational risk arising from documentation complexity, billing rules, insurance requirements, coordination needs, and provider workload.

Model Development Overview

High-level prediction pipeline

The proposed prediction pipeline would assemble data from the EHR, practice management system, scheduling platform, billing system, and payer rule sources at scheduling, check-in, or early post-visit stages. Structured fields such as visit type, payer, plan, diagnosis history, referrals, and provider schedule density would be combined with derived indicators from prior notes, coding histories, and administrative outcomes [1, 9, 18]. The model would then output an encounter-level burden risk score representing the expected likelihood that the encounter will generate disproportionate administrative work, such as coding queries, denial follow-up, authorization activity, or coordination tasks. This pipeline is intended to support operational planning rather than replace professional judgment by coding, billing, clinical documentation improvement, or care management teams.

Figure 1 presents the proposed encounter-level machine learning workflow linking administrative data sources, feature engineering domains, interpretable burden prediction, proactive workflow routing, and human governance.

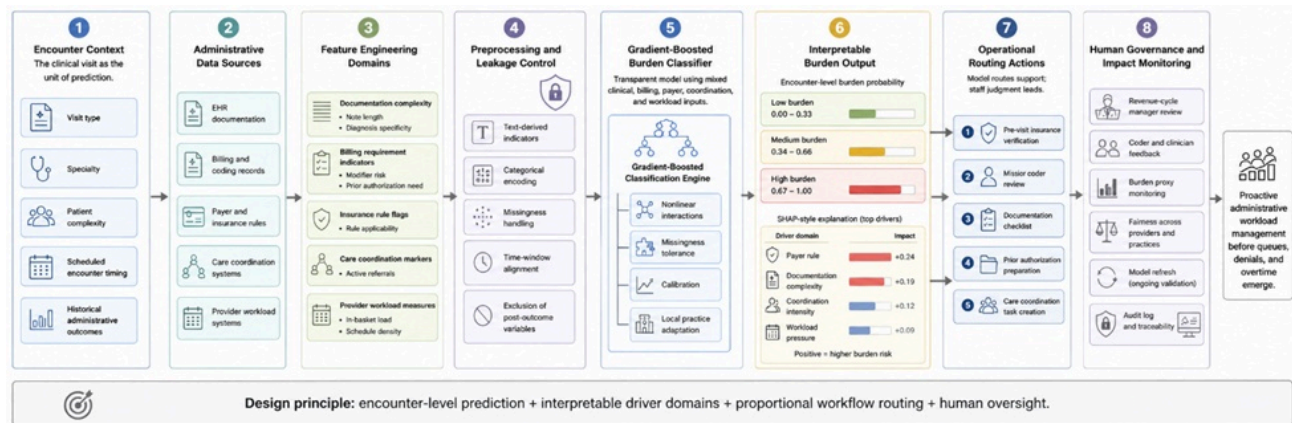


Figure 1. Machine Learning Workflow for Predicting High Administrative Burden Encounters

Core input features

The model's core input features would include a documentation complexity score, billing requirement indicators, insurance rule flags, care coordination markers, and provider workload measures. Documentation features could reflect historical note length, problem count, diagnosis specificity, and unstructured content complexity, while billing features could capture payer mix, modifier need, bundled service patterns, and historical coding intensity [8-10, 17]. Insurance and coordination features would include prior authorization flags, medical necessity checks, active referrals, recent discharge status, and specialty involvement, while workload features would include schedule density, panel characteristics, inbox queue, and documentation carryover [11-13, 21]. Combining these domains would allow the model to represent administrative burden as an interaction between what the patient encounter requires and what the local practice can absorb at that moment.

Design principles

The model should be designed for early prediction, interpretability, operational actionability, and sensitivity to local practice variation. Early prediction is important because the burden score is most useful when it can trigger pre-visit insurance verification, coding support, documentation prompts, or care coordination preparation before queues form [18, 26]. Interpretability is equally important because operations managers, coders, and clinicians need to understand whether a high-risk score is driven by documentation complexity, payer rules, care coordination needs, or workload pressure [15, 22]. Sensitivity to local variation is necessary because the same encounter may impose different administrative demands across specialties, payer contracts, staffing models, and EHR workflows [22, 27].

Data Sources and Feature Engineering

Extracting documentation complexity and billing features

Documentation complexity features would be extracted from structured and unstructured EHR elements, including problem lists, visit diagnosis counts, prior note lengths, clinical text density, and historical documentation patterns by visit type. Automated clinical coding research supports the use of clinical text and diagnosis representations to infer coding needs, while documentation studies show that note composition and EHR time can vary across specialties and physicians [9, 10, 13, 19]. Billing features would be engineered from CPT and HCPCS codes, ICD-10 specificity, modifier histories, bundled service indicators, claim edit patterns, and payer-specific coding rules when available. Together, these

features would estimate the likely administrative effort needed to transform clinical documentation into an accurate and compliant bill.

Encoding insurance rules and care coordination needs

Insurance rule features would represent prior authorization requirements, payer-specific medical necessity checks, eligibility issues, benefit restrictions, and historical denial patterns associated with similar encounters. Claim denial and insurance decision support research indicates that payer-facing outcomes can be predicted from clinical, billing, and administrative inputs, supporting the inclusion of payer rule features in a burden model [18, 21, 25]. Care coordination features would include the number of active referrals, recent discharge indicators, specialty involvement, unresolved care gaps, and the need for multidisciplinary follow-up. Because payer rules and coordination needs often interact, the model would be expected to identify cases where complex care plans require unusually intensive administrative navigation.

Provider workload and practice capacity metrics

Provider workload and practice capacity features would be time-varying measures that describe the operational environment surrounding the encounter. These could include clinician daily census, appointment density, time of day, day of week, current in-basket message counts, documentation carryover, team staffing ratios, and recent overtime signals [1, 11, 22]. Studies of EHR workload trends and usability suggest that administrative pressure is shaped not only by the clinical content of care but also by cumulative system demands and available support capacity [22, 27]. Incorporating workload features would allow the model to distinguish inherently complex encounters from encounters that become high burden because they occur during periods of limited operational slack.

Table 1 defines the encounter-level feature domains through which documentation complexity, billing requirements, payer rules, coordination needs, and provider workload become measurable predictors of administrative burden.

Table 1. Encounter-Level Feature Domains for Predicting High Administrative Burden

Feature domain	Operational meaning	Example predictors	Burden mechanism represented	Expected model signal	Primary action enabled
Documentation complexity	Captures how difficult the encounter record may be to convert into complete, specific, compliant documentation	Prior note length, diagnosis count, problem-list density, clinical text complexity, historical note revision patterns, diagnosis specificity	Complex documentation increases coding review, clarification requests, compliance review, and documentation improvement workload	Higher risk when notes are long, diagnostically dense, specialty-specific, or historically associated with coding clarification	Documentation checklist, CDI review, targeted clinician prompt, early coder review
Billing requirement intensity	Represents the administrative complexity of transforming the encounter into a compliant claim	CPT or HCPCS complexity, modifier history, bundled service indicators, coding-level variation, claim edit patterns, historical charge lag	Billing rules create rework when codes, modifiers, service combinations, or documentation support are incomplete or ambiguous	Higher risk when similar encounters have required manual coding review, edits, or charge correction	Senior coder assignment, pre-bill review, billing rule validation

Insurance and payer rule exposure	Captures payer-specific requirements that may trigger verification, authorization, or denial risk	Payer type, plan category, prior authorization flags, medical necessity rules, eligibility restrictions, historical payer denials, benefit limitations	Payer rules increase administrative burden through authorization work, documentation requests, appeals, and denial follow-up	Higher risk when plan-specific rules interact with complex diagnoses, procedures, referrals, or documentation gaps	Pre-visit eligibility check, authorization preparation, payer documentation packet
Care coordination complexity	Represents cross-service work needed after or around the encounter	Active referrals, recent hospitalization, specialty involvement, unresolved care gaps, multidisciplinary follow-up, transition-of-care status	Coordination needs create administrative work through communication, scheduling, follow-up tracking, and handoff management	Higher risk when the encounter includes multiple care pathways, recent transitions, or unresolved follow-up dependencies	Care coordinator routing, referral tracking, follow-up task creation
Provider workload and practice capacity	Describes the operational load surrounding the encounter at the time prediction is made	Schedule density, same-day visit volume, time of day, day of week, in-basket count, documentation carryover, staffing ratio, recent overtime signals	The same encounter becomes more burdensome when it occurs during limited administrative or clinical capacity	Higher risk when clinical complexity coincides with high workload, backlogged documentation, or reduced team support	Staffing adjustment, queue prioritization, temporary administrative support
Historical administrative outcome pattern	Uses prior observable outcomes to identify encounters likely to generate downstream burden	Prior coding queries, denial history, authorization follow-up, prolonged charge review, repeated payer requests, administrative overtime linked to encounter class	Past administrative outcomes reveal latent burden-generating patterns not fully captured by individual features	Higher risk when similar encounters repeatedly produce rework, delay, or unresolved administrative tasks	Risk-tier calibration, targeted work-queue routing, local model refinement
Local context and practice heterogeneity	Accounts for differences across specialty, site, staffing model, and EHR configuration	Specialty, practice site, provider panel complexity, team model, local coding practice, EHR template structure, staffing configuration	Burden varies by local workflow and should not be interpreted as a simple measure of clinician efficiency	Higher risk only after contextualizing expected baseline burden for the practice or specialty	Fair interpretation, site-specific calibration, governance review

Predictive Model Architecture

Model choice and rationale

A gradient-boosted classification model, such as an XGBoost- or LightGBM-style architecture, would be appropriate because the input space contains mixed feature types, missingness patterns, nonlinear relationships, and interactions across clinical, administrative, and workload domains. Although health system-scale language models and deep learning approaches can support broad prediction tasks, an operational burden model must remain interpretable enough for revenue cycle managers, coders, and practice leaders to act on its outputs [24, 26]. Gradient boosting can represent interactions such as payer-specific rules becoming more burdensome when documentation complexity and provider workload are simultaneously elevated. The model would be positioned as a decision-support tool rather than an autonomous administrative decision-maker.

Input feature vector and preprocessing

The encounter-level input feature vector would combine continuous variables, categorical variables, binary indicators, historical aggregates, and time-sensitive operational measures. Continuous features such as prior note length, diagnosis count, schedule density, and in-basket volume would be normalized or transformed as needed, while payer, plan, specialty, visit type, and coding category would be encoded to preserve operational meaning [15, 17, 19]. Missing workload data would require explicit handling because absent in-basket or staffing information may reflect measurement limitations rather than true low burden. Preprocessing would also need to prevent leakage by ensuring that features available only after denials, coding queries, or overtime events are not used for predictions intended at scheduling or check-in.

Output: encounter-level burden probability

The model output would be a calibrated encounter-level probability that the visit will generate high administrative burden, with the score optionally translated into low, medium, or high tiers for operational use. The target outcome would be conceptually defined using observable proxy signals such as excess coding queries, payer documentation requests, claim denial activity, prolonged charge review, repeated prior authorization follow-up, or administrative staff overtime linked to the encounter [18, 21, 20]. Because administrative burden is not directly measured in most systems, the outcome definition should be transparent, locally validated, and refined with input from coding, billing, clinical documentation improvement, and care coordination teams. The output should guide proportional support rather than create punitive comparisons across clinicians or practices.

Handling Temporal Dynamics and Provider Variability

Time-varying workload and seasonal effects

Administrative burden risk is likely to fluctuate over time because schedule density, documentation carryover, in-basket queues, staffing shortages, and payer processing cycles change from day to day. The proposed model would incorporate time-varying indicators such as same-day appointment load, recent message volume, recent documentation after-hours work, and local administrative queue status to capture transient capacity strain [1, 2, 11]. Seasonal effects could also be represented through calendar features, payer renewal periods, coding update cycles, and predictable surges in authorization or eligibility work. These temporal features would allow the model to distinguish stable encounter complexity from temporary operational congestion.

Accounting for provider and practice heterogeneity

Provider and practice heterogeneity must be handled carefully so the model does not misclassify high-complexity practices or high-documentation specialties as inefficient. Provider-specific features, specialty indicators, practice-site

variables, and historical documentation patterns could help contextualize burden predictions while preserving the encounter-level focus [12–14]. Random intercepts or locally calibrated feature effects could represent baseline differences in patient panel complexity, staffing models, EHR configuration, and specialty-specific documentation requirements. This approach would support fairer operational interpretation by separating predictable local workflow variation from truly exceptional administrative burden.

Continuous learning and model refreshing

Continuous learning is necessary because administrative burden is shaped by changing payer policies, coding rules, documentation standards, clinical workflows, and staffing models. Periodic model refreshing would allow the system to adapt when new prior authorization rules, claim edits, documentation templates, or coding guidance alter the relationship between encounter features and downstream administrative work [18, 21, 27]. Refresh cycles should be governed rather than automatic, with monitoring for shifts in feature distributions, proxy outcome definitions, and operational use patterns. Model updates should therefore combine statistical recalibration with review by revenue cycle, clinical documentation improvement, and care coordination stakeholders.

Model Interpretability and Decision Support

Explaining burden predictions to operations managers

Interpretability is essential because administrative teams need to understand why a specific encounter is predicted to create high burden. Feature attribution methods, such as SHAP-style explanations, could identify whether the risk is driven by documentation complexity, a payer-specific authorization rule, multiple active referrals, recent discharge status, or a provider workload surge [15, 19, 26]. For example, an operations manager could see that a high predicted burden reflects the combination of complex diagnoses, a payer documentation requirement, and a dense clinician schedule rather than a generic high-risk label. Such explanations would make the model more actionable and reduce the risk that predictions are interpreted as opaque judgments about clinicians.

Proactive resource allocation and workflow routing

Decision support should translate burden predictions into operational actions that reduce downstream rework. High-risk encounters could be routed for pre-visit insurance verification, senior coder review, documentation improvement prompts, authorization preparation, or care coordination outreach before the claim or follow-up task becomes delayed [18, 21, 26]. Automated documentation support and coding assistance may be especially useful when the predicted burden is driven by incomplete specificity, complex ICD mapping, or payer-sensitive coding requirements [9, 16, 28]. The model should therefore be embedded in workflow logic that recommends proportionate support rather than simply displaying a risk score.

Operational Integration and Proactive Workload Management

Embedding in scheduling and revenue cycle systems

Operational integration would require the burden score to appear in systems already used by scheduling, coding, billing, clinical documentation improvement, and care coordination teams. Displaying the score on daily schedules or work queues would allow managers to pre-assign support, adjust coder coverage, or coordinate payer verification before a high-risk encounter reaches the post-visit revenue cycle stage [18, 23, 25]. Integration with EHR and practice management workflows would also help align prediction timing with the moment when staff can still intervene. The goal is not to add another dashboard but to place the prediction where administrative decisions are already made.

Triggering automated process steps

A high burden score could trigger automated process steps such as eligibility verification, prior authorization screening, documentation checklist generation, coder notification, or care coordination task creation. These steps would be most useful when tied to the specific feature drivers behind the prediction, such as a payer rule, a complex diagnosis cluster, or an unusually heavy provider workload context [20, 21, 26]. Automated coding and documentation support tools suggest that administrative interventions can be embedded into clinical and billing workflows, but they must be designed to reduce work rather than create additional prompts or alert fatigue [15, 16, 28]. The proposed model would therefore function as a routing layer that activates the right support only when predicted burden justifies it.

Table 2 translates predicted burden drivers into proportional administrative actions, responsible human reviewers, and measurable operational impact signals.

Table 2. Prediction-to-Action Framework for High Administrative Burden Encounter Management

Predicted burden driver	Example explanation shown to users	Operational interpretation	Recommended proportional action	Human reviewer	Evaluation signal
Documentation complexity dominant	“Risk is mainly driven by diagnosis specificity, long prior notes, and historically high coding clarification for similar visits.”	The encounter may require additional documentation specificity before billing or coding can proceed efficiently	Generate documentation checklist; route to CDI support; flag likely coding clarification needs	Clinical documentation improvement specialist; coder; clinician reviewer when needed	Reduction in documentation queries, note revision cycles, and coding clarification delay
Billing requirement dominant	“Risk is mainly driven by modifier history, bundled service pattern, and coding-level ambiguity.”	The encounter is likely to require manual billing review or coding expertise	Assign senior coder; perform pre-bill validation; review modifier and bundled-service rules	Coding supervisor or revenue-cycle manager	Reduction in claim edits, charge lag, and manual rework
Insurance rule dominant	“Risk is mainly driven by payer-specific authorization and medical necessity requirements.”	The encounter may generate payer documentation, eligibility, authorization, or denial-related work	Trigger eligibility verification; prepare prior authorization packet; check medical necessity documentation	Prior authorization team; payer specialist	Reduction in authorization delay, payer documentation requests, and avoidable denials
Care coordination dominant	“Risk is mainly driven by recent discharge status, multiple active referrals, and unresolved care gaps.”	Administrative burden is likely to arise from communication, scheduling, and cross-team follow-up	Create care coordination task; prioritize referral follow-up; assign navigator or coordinator	Care coordinator; practice manager	Reduction in unresolved referrals, delayed follow-up, and repeated coordination messages
Provider workload	“Risk is mainly driven by dense	The encounter may become high	Add temporary administrative	Practice manager;	Reduction in staff overtime, after-

dominant	clinic schedule, high in-basket volume, and documentation carryover.”	burden because local capacity is limited even if the clinical case is not unusually complex	support; shift coding queue priority; adjust same-day task assignment	operations lead	hours documentation, and accumulated administrative queue
Multidomain high risk	“Risk reflects combined payer, documentation, coordination, and workload pressure.”	The encounter is likely to produce cumulative administrative work across multiple teams	Bundle actions: payer check, coder review, documentation checklist, and coordinator routing	Revenue-cycle manager plus clinical operations lead	Reduction in total administrative touchpoints, denial-related rework, and queue escalation
Low predicted burden with stable explanation	“No major payer, documentation, coordination, or workload driver is elevated.”	Routine processing is likely sufficient; intervention may create unnecessary work	No additional routing; standard work queue handling	Routine staff oversight only	Maintenance of low false-negative rate and avoidance of alert fatigue
High uncertainty or missing data	“Prediction is uncertain because workload, payer rule, or documentation indicators are incomplete.”	The model lacks enough reliable information for confident routing	Send to human triage rather than automated action; improve data completeness	Operations reviewer; data governance lead	Monitoring of missingness, override frequency, and uncertain-risk outcomes

Evaluation Strategy

Predictive performance metrics

The model should be evaluated using classification, calibration, and decision-support criteria appropriate for high-burden encounter prediction. AUROC, precision-recall analysis, and calibration assessment could be used conceptually to determine whether the model distinguishes high-burden encounters from routine encounters and whether predicted probabilities are meaningful for operational triage [18, 21, 23]. Comparisons with a baseline model using only encounter type or payer category would clarify whether documentation, coordination, and workload features add useful signal. Evaluation should avoid focusing only on discrimination because an operational model must also produce stable, explainable, and actionable risk tiers.

Temporal and external validation

Temporal validation would test whether the model remains useful when applied to future encounters after payer rules, staffing patterns, coding practices, or documentation workflows have shifted. Forward-time testing is especially important because EHR workload, note composition, and administrative processes can change across implementation periods and health system contexts [13, 22, 27]. External validation at another clinic, specialty, or health system would assess whether the feature domains are portable or whether local calibration is required. Such validation should emphasize conceptual robustness, workflow fit, and governance readiness rather than assuming that one model specification will generalize unchanged.

Operational impact

Operational impact should be evaluated by examining whether model-guided workflows reduce preventable coding queries, denial-related rework, authorization delays, documentation clarification cycles, and administrative overtime. These outcomes align with prior work on claim denial prediction, documentation burden measurement, administrative waste, and clinician workload, but they should be interpreted as operational signals rather than direct proof of reduced burden for every stakeholder [3, 5, 18, 20]. The evaluation should also examine whether proactive routing changes staff workload distribution or unintentionally shifts burden from billing teams to clinicians. A successful implementation would be expected to improve coordination between revenue cycle and clinical operations while preserving fairness and usability.

Limitations

Data availability and labeling

A major limitation is that administrative burden is rarely measured directly at the encounter level, so the model would depend on proxy labels such as denial activity, coding queries, charge lag, authorization follow-up, or staff overtime. These proxies may capture only visible administrative outcomes while missing informal work performed by clinicians, care coordinators, or billing staff outside structured systems [1, 11, 20]. Documentation burden measurement research underscores the difficulty of translating complex work into reliable operational indicators, especially when EHR logs, note metadata, and staff activity systems capture different parts of the process. The proposed model would therefore require careful local label design and ongoing review to ensure that predictions reflect meaningful administrative burden rather than convenience of measurement.

Adaptability to changing payer policies

A second limitation is that payer policies, medical necessity criteria, authorization requirements, and coding edits may change faster than model retraining cycles. When these rules shift, a previously reliable feature pattern may no longer predict the same level of administrative work, particularly for specialties or payers with frequent documentation changes [18, 21, 25]. Continuous monitoring and periodic refreshes can reduce this risk, but they do not eliminate the need for human governance and payer-policy surveillance. The model should therefore be treated as an adaptive operational aid rather than a static representation of administrative complexity.

Conclusion

A machine learning model for predicting high administrative burden encounters could provide an encounter-level risk signal before administrative work becomes visible as a backlog, denial, coding query, or overtime event. By combining documentation complexity, billing requirements, insurance rules, care coordination needs, and provider workload indicators, the model would frame burden as a multidomain operational risk. This approach would allow healthcare organizations to anticipate which encounters are most likely to require extra administrative support.

The main strength of the proposed model is its early and integrated design. It would use information available around scheduling, check-in, or early post-visit processing to support timely intervention rather than delayed remediation. Interpretable outputs would help managers understand whether risk is driven by payer rules, documentation intensity, coordination complexity, or workload strain.

Important challenges remain before such a model could be implemented responsibly. Administrative burden must be defined through imperfect proxy outcomes, and those outcomes may vary by health system, specialty, staffing model, and payer environment. Data freshness, model governance, local calibration, and workflow integration would determine whether the model reduces administrative work or merely reorganizes it.

Future pilot studies should examine this approach in large multi-specialty practices where encounter complexity, payer variation, and administrative queues are substantial. Such pilots should assess whether burden prediction improves time allocation, revenue integrity, documentation quality, and coordination workflows. The broader goal is to move healthcare administration toward proactive support systems that reduce avoidable rework while preserving clinician and staff capacity.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 01 Feb 2026

Revised: 08 Mar 2026

Accepted: 12 Apr 2026

Published online: 20 July 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

- Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan WJ, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med*. 2017;15(5):419-26.
- Tai-Seale M, Olson CW, Li J, Chan AS, Morikawa C, Durbin M, et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. *Health Aff*. 2017;36(4):655-62.
- Shrank WH, Rogstad TL, Parekh N. Waste in the US health care system: estimated costs and potential for savings. *JAMA*. 2019;322(15):1501-9.
- Kocher RP. Reducing administrative waste in the US health care system. *JAMA*. 2021;325(5):427-8.
- Chernew M, Mintz H. Administrative expenses in the US health care system: why so high? *JAMA*. 2021;326(17):1679-80.

Melnick ER, Dyrbye LN, Sinsky CA, Trockel M, West CP, Nedelec L, et al. The association between perceived electronic health record usability and professional burnout among US physicians. *Mayo Clin Proc.* 2020;95(3):476-87.

Harry E, Sinsky C, Dyrbye LN, Makowski MS, Trockel M, Tutty M, et al. Physician task load and the risk of burnout among US physicians in a national survey. *Jt Comm J Qual Patient Saf.* 2021;47(2):76-85.

Zheng J, Yu H. Assessing the readability of medical documents: a ranking approach. *JMIR Med Inform.* 2018;6(1):e8611.

Catling F, Spithourakis GP, Riedel S. Towards automated clinical coding. *Int J Med Inform.* 2018;120:50-61.

Atutxa A, de Ilaraza AD, Gojenola K, Oronoz M, Perez-de-Viñaspre O. Interpretable deep learning to map diagnostic texts to ICD-10 codes. *Int J Med Inform.* 2019;129:49-59.

Lou SS, Lew D, Harford DR, Lu C, Evanoff BA, Duncan JG, et al. Temporal associations between EHR-derived workload, burnout, and errors: a prospective cohort study. *J Gen Intern Med.* 2022;37(9):2165-72.

Rittenberg E, Liebman JB, Rexrode KM. Primary care physician gender and electronic health record workload. *J Gen Intern Med.* 2022;37(13):3295-301.

Rotenstein LS, Apathy N, Holmgren AJ, Bates DW. Physician note composition patterns and time on the EHR across specialty types: a national, cross-sectional study. *J Gen Intern Med.* 2023;38(5):1119-26.

Goldstein IH, Hwang T, Gowrisankaran S, Bales R, Chiang MF, Hribar MR. Changes in electronic health record use time and documentation over the course of a decade. *Ophthalmology.* 2019;126(6):783-91.

Dong H, Falis M, Whiteley W, Alex B, Matterson J, Ji S, et al. Automated clinical coding: what, why, and where we are? *NPJ Digit Med.* 2022;5:159.

Sonabend A, Cai W, Ahuja Y, Ananthakrishnan A, Xia Z, Yu S, et al. Automated ICD coding via unsupervised knowledge integration (UNITE). *Int J Med Inform.* 2020;139:104135.

Moons E, Khanna A, Akkasi A, Moens MF. A comparison of deep learning methods for ICD coding of clinical records. *Appl Sci.* 2020;10(15):5262.

Johnson M, Albizri A, Harfouche A. Responsible artificial intelligence in healthcare: predicting and preventing insurance claim denials for economic and social wellbeing. *Inf Syst Front.* 2023;25(6):2179-95.

Diao X, Huo Y, Zhao S, Yuan J, Cui M, Wang Y, et al. Automated ICD coding for primary diagnosis via clinically interpretable machine learning. *Int J Med Inform.* 2021;153:104543.

Murad MH, Stelling BE, West CP, Hasan B, Simha S, Saadi S, et al. Measuring documentation burden in healthcare. *J Gen Intern Med.* 2024;39(14):2837.

Su SC, Huang CC, Gung RR, Hsiung LK, Gao ZW, Tsai CE. A decision support system with artificial intelligence and natural language processing to mitigate the deduction rate of health insurance claims. *Appl Sci.* 2021;11(24):11623.

Holmgren AJ, Hendrix N, Maisel N, Everson J, Bazemore A, Rotenstein L, et al. Electronic health record usability, satisfaction, and burnout for family physicians. *JAMA Netw Open.* 2024;7(8):e2426956.

Saripalli P, Tirumala V, Chimmad A. Assessment of healthcare claims rejection risk using machine learning. In: *IEEE Healthcom.* 2017. p. 1-6.

Jiang LY, Liu XC, Nejatian NP, Nasir-Moin M, Wang D, Abidin A, et al. Health system-scale language models are all-purpose prediction engines. *Nature.* 2023;619(7969):357-62.

Owolabi T. Transforming appeal decisions: machine learning triage for hospital admission denials. *JAMIA Open*. 2025;8(1):ooaf016.

Perkins SW, Muste JC, Alam T, Singh RP. Improving clinical documentation with artificial intelligence: a systematic review. *Perspect Health Inf Manag*. 2024;21(2).

Arndt BG, Micek MA, Rule A, Shafer CM, Baltus JJ, Sinsky CA. More tethered to the EHR: EHR workload trends among academic primary care physicians, 2019-2023. *Ann Fam Med*. 2024;22(1):12-8.

Genes N, Sills J, Heaton HA, Shy BD, Scofi J. Addressing note bloat: solutions for effective clinical documentation. *JACEP Open*. 2025;6(1):100031.