

RESEARCH ARTICLE

Interpretable AI-Enabled Framework for Proactive Revenue Cycle Management: An Evidence-Informed Integrative Review

Remon Gameel Ebed Saweris ¹

¹*Institute of Military Health and Preventive Medicine, Military Medical Academy, Cairo, Egypt*

*Correspondence remongameel@gmail.com

Received: 13 April 2026; Revised: 28 May 2026; Accepted: 29 September 2026, Published: 30 September 2026

ABSTRACT

Revenue cycle management (RCM) is a data-intensive healthcare process that connects patient access, authorization, clinical documentation, coding, claims processing, denial management, and payment. Although artificial intelligence (AI) is increasingly applied across these activities, the available evidence remains fragmented across administrative, clinical, and financial workflows, limiting its translation into an integrated and accountable RCM strategy. This study develops an interpretable AI-enabled framework for proactive RCM through an evidence-informed integrative review. The review synthesized evidence from administrative health records, claims analytics, automated International Classification of Diseases coding, AI-assisted documentation, prior authorization, denial and appeal triage, explainable AI, and trustworthy healthcare AI. The evidence base incorporated systematic reviews covering 70 administrative-health-record studies, 137 healthcare-claims fraud studies, 16 prior-authorization studies, 73 automated coding studies using Medical Information Mart for Intensive Care data, and 519 healthcare large-language-model evaluations. The synthesis shows that AI applications are more established for documentation and coding assistance, risk scoring, denial prioritization, and anomaly detection, whereas direct evidence for end-to-end financial improvement in RCM remains limited. The proposed framework integrates four layers: data and interoperability, predictive intelligence, workflow orchestration, and governance with human oversight. The framework also links predicted risk, financial impact, and intervention cost to decision prioritization. This study provides a structured architecture for human-led, AI-augmented RCM and defines technical, operational, financial, and governance outcomes for future prospective evaluation.

KEYWORDS

artificial intelligence; machine learning; revenue cycle management; health informatics; medical coding; claims management; denial prediction; clinical documentation; explainable AI; healthcare finance

1. Introduction

Revenue cycle management (RCM) connects healthcare services with administrative and financial processes that support reimbursement. The process begins with patient registration, eligibility verification, and authorization. It continues with clinical documentation, coding, charge capture, claim submission, adjudication,

payment posting, denial resolution, and accounts-receivable follow-up. Errors in one stage can affect subsequent stages, increase administrative work, delay reimbursement, and create revenue loss.

Healthcare administrative data contain structured information about patient encounters, diagnoses, procedures, payers, utilization, adjudication, and payment. These data support the development of machine-learning models for healthcare administration. A systematic review identified 70 studies that applied machine learning to administrative health records and reported substantial variation in data sources and analytical methods [1]. Claims-based machine-learning studies also used large administrative datasets to predict healthcare outcomes and included independent validation in several settings [2, 3].

Healthcare organizations have also started to use artificial intelligence (AI) in revenue-cycle activities. A 2024 survey from the American Hospital Association reported that 46% of participating hospitals and health systems used AI in RCM-related operations [4]. AI applications also support clinical documentation and coding. The U.S. Government Accountability Office reported the use of AI tools for documentation and coding and identified issues related to accuracy, verification, and oversight [5].

The available evidence remains distributed across different administrative and clinical functions. Automated coding studies often focus on classification performance. Claims-based studies frequently evaluate clinical or utilization outcomes. Documentation studies commonly assess documentation time and quality. These studies do not consistently evaluate denial prevention, reimbursement, or financial outcomes within an integrated RCM workflow. This condition creates an implementation gap for healthcare organizations. A predictive model does not improve an RCM process unless its output supports a specific operational action. An AI-enabled RCM system therefore requires data quality, interoperability, explainability, external validation, privacy protection, auditability, workflow integration, and human accountability.

RCM also requires a clear separation between prediction and action. An AI model can estimate denial risk, identify documentation gaps, recommend billing codes, or prioritize appeals. Healthcare staff must interpret these outputs before taking consequential administrative or financial actions. The decision process must consider operational workload, financial impact, compliance requirements, and the consequences of incorrect recommendations.

This study addresses this gap through an evidence-informed integrative review. The study synthesizes evidence from administrative health records, claims analytics, automated coding, AI-assisted documentation, prior authorization, denial and appeal management, explainable AI, and trustworthy healthcare AI. The proposed framework organizes these capabilities into four layers: data and interoperability, predictive intelligence, workflow orchestration, and governance with human oversight.

This study has four objectives. First, the study maps AI capabilities to principal RCM decision points. Second, the study synthesizes evidence related to administrative and financial healthcare workflows. Third, the study identifies implementation and governance requirements for AI-enabled RCM. Fourth, the study develops an interpretable human-in-the-loop framework for proactive RCM.

This study provides four contributions. First, the study connects AI capabilities with specific RCM decision points. Second, the study proposes a four-layer framework that integrates data, predictive intelligence, workflow orchestration, and governance. Third, the study separates technical model performance from operational and financial outcomes. Fourth, the study defines a research agenda that emphasizes external validation, calibration, human oversight, and measurable RCM outcomes. The study provides a conceptual framework based on published evidence and does not train a new predictive model or analyze a new patient-level dataset.

2. Preliminaries: RCM Decision and Risk Modeling

This study uses an evidence-informed integrative review to examine the application of artificial intelligence in revenue cycle management (RCM). The review covers administrative health records, claims analytics, clinical documentation, coding, prior authorization, denial management, and healthcare AI governance. The study does not calculate a pooled effect because the reviewed evidence uses different data sources, methods, outcomes, and implementation settings.

The conceptual model represents each claim through administrative, clinical, coding, documentation, payer, and workflow features. Let $\mathbf{x}_i$ denote the feature vector for claim i , and let D_i denote the denial outcome. The predicted denial probability is defined as

$$p_i = P(D_i = 1 \mid \mathbf{x}_i), \quad (1)$$

where p_i represents the estimated probability that claim i will experience the target denial event.

The framework also represents the financial value of an intervention. Let V_i denote the financial impact associated with claim i , and let C_i denote the expected intervention cost. The expected intervention value is defined as

$$EV_i = P(\text{Recoverable}_i \mid \mathbf{x}_i) \times V_i - C_i, \quad (2)$$

where $P(\text{Recoverable}_i \mid \mathbf{x}_i)$ represents the estimated probability that the intervention will recover or protect the financial value associated with claim i .

The formulation separates prediction from operational decision-making. The predictive component estimates denial risk or recoverability. The decision component combines the predicted outcome with financial impact and intervention cost. This structure separates model output from the final action taken by healthcare staff. Table 1 summarizes the variables used in the RCM decision formulation.

Table 1. Conceptual variables used in the RCM risk and expected-value formulation.

Symbol	Definition	RCM Interpretation
$\mathbf{x}_i$	Feature vector for claim i	Patient, payer, provider, coding, documentation, and workflow features
D_i	Binary denial outcome	1 if the target denial occurs; 0 otherwise
p_i	Predicted denial probability	Estimated denial risk before claim submission
V_i	Financial impact	Expected recoverable or exposed financial value associated with claim i
C_i	Intervention cost	Expected staff or system cost of acting on claim i
EV_i	Expected intervention value	Expected recoverable value minus intervention cost

The equations serve as conceptual decision specifications. The study does not train a production model and does not estimate patient-level parameters. The formulation provides a consistent structure for linking predictive outputs with operational and financial decisions in RCM.

3. Methodology

3.1. Structured Evidence Search and Evidence Eligibility

This study uses a structured and targeted evidence-search strategy to identify literature related to artificial intelligence in RCM. The search covers artificial intelligence, machine learning, natural language processing, large language models, administrative health records, claims, coding, documentation, denial, reimbursement, and revenue cycle management. The search prioritizes peer-reviewed systematic reviews and primary studies published from 2020 onward. The study also includes foundational methodological studies and authoritative guidance when they provide relevant information on validation, explainability, governance, or healthcare AI implementation. Studies are included when they provide direct evidence for one or more RCM decision points. Studies are also included when they provide transferable evidence related to claims prediction, coding, documentation, validation, explainability, or healthcare AI governance.

The study uses an evidence-informed integrative review rather than a registered systematic review. The review does not claim exhaustive literature retrieval. The study does not report a PRISMA screening count because it does not use a registered systematic-review design. The study reports quantitative study counts only when the underlying reviews or primary studies provide those values. Figure 1 presents the evidence-review and framework-construction process. The process starts with the definition of search concepts and eligibility criteria. The process continues with evidence retrieval, relevance screening, evidence extraction, RCM-domain mapping, assessment of implementation and governance implications, framework construction, and identification of evidence gaps and validation requirements.

3.2. Evidence Extraction and Integrative Synthesis

The study records the extracted evidence in a structured conceptual matrix. The matrix contains the RCM domain, AI capability, data source, supported decision, reported outcome, validation approach, implementation constraint, and governance implication. The synthesis groups evidence according to similarities across studies. The study does not combine heterogeneous effect estimates into a pooled value. Evidence from related domains is classified as transferable evidence when it supports an RCM decision, implementation requirement, or

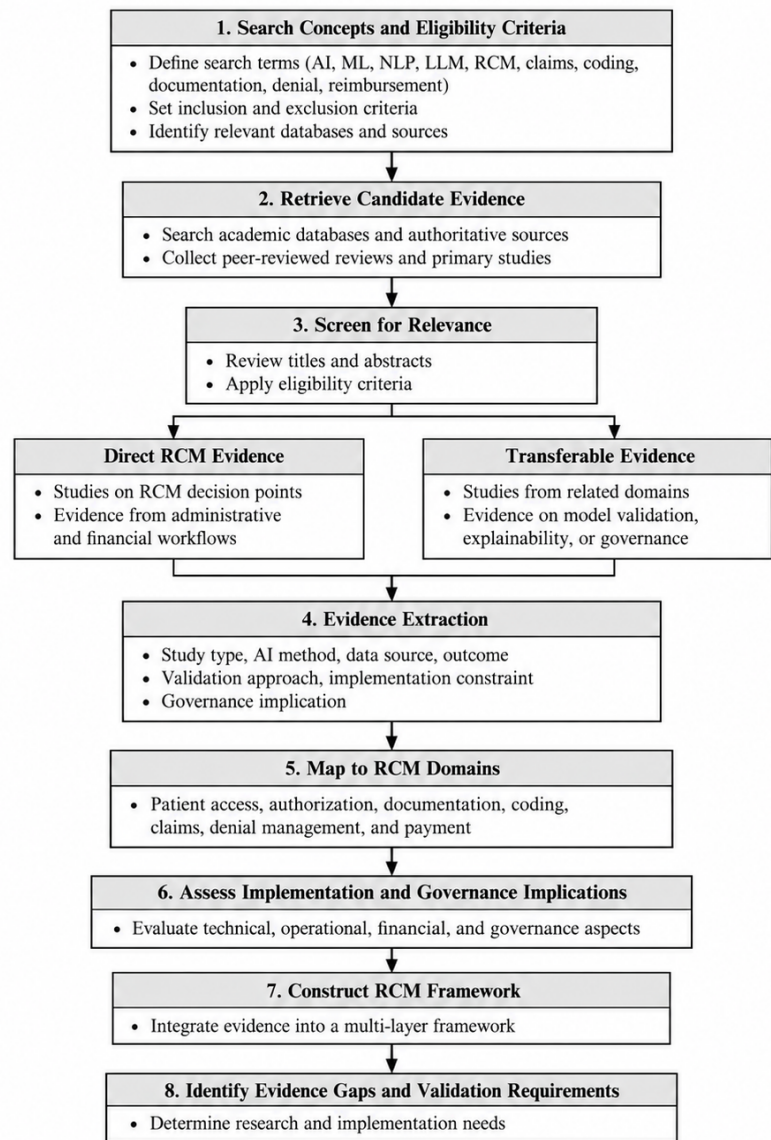


Figure 1. Workflow of the evidence-informed integrative review and framework construction.

governance principle. The synthesis separates technical model performance from operational RCM outcomes. A model can achieve high predictive performance without demonstrating a reduction in denials, administrative workload, or financial loss. This distinction guides the interpretation of the reviewed evidence.

3.3. Evidence Mapping Across RCM Domains

The study organizes the evidence into six RCM domains. The domains include patient access and eligibility, authorization and coverage, documentation and coding, claim quality and denial prevention, denial and appeal management, and revenue integrity and anomaly detection. The study evaluates each domain based on its AI capability, data requirements, supported decision, operational value, implementation risk, and available evidence. This mapping connects technical methods with specific RCM workflows.

3.4. Framework Construction and Decision Formulation

The framework links each RCM domain with its decision point, data requirement, AI capability, human intervention, and measurable outcome. The framework separates predictive output from operational action. For claim i , the predicted denial probability is defined as

$$p_i = P(D_i = 1 \mid \mathbf{x}_i), \quad (3)$$

where $\mathbf{x}_i$ represents the feature vector for claim i , and D_i represents the denial outcome. The framework defines the expected intervention value as

$$EV_i = P(\text{Recoverable}_i \mid \mathbf{x}_i) \times V_i - C_i, \quad (4)$$

where V_i represents the financial impact and C_i represents the intervention cost. The equations serve as conceptual decision specifications. The study does not train a production model. The study also does not estimate claim-level model parameters. The formulation connects predicted risk with operational and financial decision criteria.

3.5. Reproducible Review Workflow

The review follows a defined sequence. The process begins by defining search concepts and eligibility criteria. The process then retrieves candidate evidence and screens the evidence for relevance. The study extracts the selected evidence and maps it to the relevant RCM domains. The study then evaluates implementation and governance implications, constructs the framework, and identifies evidence gaps and validation requirements. The review produces an evidence-informed RCM framework and a research agenda for future prospective evaluation.

4. Results and Discussion

4.1. Patient Access, Eligibility and Authorization

Patient access, eligibility verification, and authorization form the initial stages of revenue cycle management. Errors at these stages can affect claim preparation and subsequent administrative processes. Artificial intelligence can process information related to patients, payers, services, providers, and previous transactions. This information can support the identification of missing eligibility requirements, incomplete authorization, and inconsistent administrative data before claim submission. Published studies on artificial intelligence for prior authorization mainly report technical performance metrics. These metrics include precision, recall, F1-score, and area under the receiver operating characteristic curve. Fewer studies evaluate operational outcomes such as denial reduction, processing cost, and patient-access improvement [6]. This evidence supports a risk-based routing approach for authorization workflows. The system can process routine cases with complete information through predefined rules. The system can route uncertain cases to staff for review. Cases with greater financial or administrative consequences can receive additional human verification. Artificial intelligence should therefore support case prioritization and workflow routing rather than replace human review. Human staff remain responsible for decisions related to coverage, authorization, and claim eligibility.

4.2. Documentation and Coding

Clinical documentation and coding have a direct relationship with claim preparation and reimbursement. Artificial intelligence can support these activities by extracting information from clinical notes, identifying missing details, suggesting codes, and detecting inconsistencies before claim submission. Recent studies report the use of artificial intelligence to support documentation efficiency and coding processes. A systematic review of AI-powered documentation systems reported potential improvements in documentation efficiency but also identified variation in documentation quality and concerns about reliability [7]. A scoping review also reported improvements in documentation efficiency and accuracy across different clinical settings, together with implementation challenges [8]. A systematic review and meta-analysis found a reduction in documentation workload, but the included studies showed substantial heterogeneity and generally low methodological quality [9].

Artificial intelligence has also been applied to automated clinical coding. A systematic review of AI-based International Classification of Diseases coding identified benefits related to efficiency, standardization, and coding accuracy, while also reporting challenges related to large label spaces, class imbalance, long clinical documents, interpretability, and transparency [10]. Another review described the development of automated coding methods from convolutional and recurrent architectures toward attention mechanisms, transformers, and pretrained language models [11]. A review based on the Medical Information Mart for Intensive Care literature also identified limitations related to dataset diversity, clinical validation, and explainability [12]. The reviewed evidence supports the use of artificial intelligence as an assistance layer for documentation and coding. Human review remains necessary because coding decisions affect reimbursement, compliance, and claim submission. The workflow should therefore preserve a traceable connection between source documentation, AI-generated recommendations, reviewer decisions, and the final claim.

4.3. Claim-Quality Prediction and Denial Prevention

Claim-quality prediction supports the identification of potential denial risk before claim submission. Historical claims and remittance data can provide information about patterns associated with previous denials. For claim i , the denial risk can be represented as

$$\text{DenialRisk}_i = P(D_i = 1 \mid X_i),$$

Previous studies using administrative claims data show that machine-learning models can support prediction tasks based on large healthcare datasets [2, 3]. These studies provide evidence for the predictive use of administrative data, but they do not directly establish improvements in denial prevention or revenue outcomes. The use of denial-risk prediction should therefore focus on case prioritization before claim submission. The predicted risk can support the identification of claims that require additional documentation, coding review, authorization verification, or staff assessment.

4.4. Denial Triage and Appeals

Denial management requires prioritization because each denied claim has a different probability of recovery and a different financial value. Artificial intelligence can support this process by estimating the expected recoverable value of each denied claim. A recent study applied machine-learning-based triage to support appeal decisions for hospital admission denials [13]. The study demonstrates the use of predictive classification for prioritizing cases that require review. This application provides a relevant example for denial management in revenue cycle workflows. An appeal workflow should use a ranked queue rather than a simple binary decision. The prioritization process can consider expected recoverable value, filing deadlines, clinical complexity, payer behavior, documentation availability, and previous appeal outcomes. The ranking process can support staff allocation by placing higher-priority cases earlier in the review queue. Human staff remain responsible for the final appeal decision and the submission of supporting documentation.

4.5. Revenue Integrity and Anomaly Detection

Revenue integrity requires consistency between healthcare services, clinical documentation, coding, charging, claims, and payment. Artificial intelligence can support this process by identifying unusual patterns in administrative and financial data. Anomaly detection can identify services that are documented but not

charged, repeated coding irregularities, unexpected payer-specific underpayments, and changes in provider-level patterns. These findings can support further review by revenue-cycle staff. A systematic review of machine learning for healthcare-claims fraud detection analyzed 137 studies and reported the use of supervised, unsupervised, and hybrid methods [14]. The review also identified challenges related to data inconsistency, limited standardization, privacy, and the availability of labeled fraud cases. Fraud detection and revenue integrity address different operational objectives. However, both areas use similar analytical methods for identifying unusual patterns in large administrative datasets. The methodological evidence from fraud-detection studies can therefore inform anomaly-detection approaches in revenue integrity. The objective of anomaly detection in revenue integrity is not to maximize the number of detected anomalies. The process should prioritize anomalies that have operational relevance, financial impact, and sufficient information for professional review. The proposed framework integrates the reviewed evidence into four layers. Table 2 summarizes these layers.

Table 2. Four-layer interpretable AI-enabled RCM framework.

Layer	Inputs / Methods	Controls	Outputs
1. Data and Interoperability	EHR/HIS, claims, payer, authorization, coding, and remittance data	Data quality, provenance, and privacy	Trusted RCM data
2. Predictive Intelligence	Machine learning, natural language processing, large language models, risk scoring, and anomaly detection	Calibration, bias assessment, external validation, and explainable AI	Risk scores and recommendations
3. Workflow Orchestration	Business rules, staff capacity, and AI outputs	Decision thresholds, human approval, and audit trails	Work queues, alerts, and drafts
4. Governance and Human Oversight	Model outputs, feedback, and operational outcomes	Security, accountability, and monitoring	Controlled system improvement

4.6. Decision Layer and Expected Value

The decision layer converts predictive outputs into operational priorities. The framework does not use predicted risk as the only basis for action. The decision process also considers financial impact and intervention cost. For claim i , the expected intervention value is defined as

$$\text{ExpectedValue}_i = P(\text{Recoverable}_i \mid X_i) \times \text{FinancialImpact}_i - \text{InterventionCost}_i, \quad (6)$$

where $P(\text{Recoverable}_i \mid X_i)$ represents the estimated probability of financial recovery, FinancialImpact_i represents the financial value associated with the claim, and $\text{InterventionCost}_i$ represents the expected cost of intervention. The formulation supports claim prioritization based on expected operational value. A claim with a higher predicted risk does not automatically receive the highest priority. The decision also depends on the expected financial impact and the cost of intervention. The framework separates prediction from action. The predictive component estimates risk or recoverability. The workflow component determines the appropriate intervention. Human staff remain responsible for decisions that involve compliance, clinical documentation, or material financial consequences. This structure allows the system to rank cases without assigning full decision authority to the predictive model. The ranked output can support work queues, review priorities, and intervention planning within the revenue cycle workflow.

4.7. Human-in-the-Loop Design

The proposed framework places human oversight at key decision points in the revenue cycle workflow. Artificial intelligence provides recommendations, predictions, or draft outputs, while authorized staff retain responsibility for final decisions. Artificial intelligence can identify missing information in clinical or administrative records. Qualified staff determine whether the identified information is required before the claim proceeds. Artificial intelligence can also recommend clinical codes. Authorized coding professionals review the recommendation and determine the final code according to organizational policy. In denial management, artificial intelligence can rank cases according to predicted recoverability or operational priority. Revenue-cycle staff review the ranked cases before deciding whether an appeal should be submitted. Artificial intelligence can also generate draft communications, documentation queries, or appeal materials. The final submission remains subject to human review and approval. The framework also requires traceable AI outputs. Explanations should identify the principal factors that contribute to a recommendation and should refer to the

relevant source information when available [15]. Human oversight separates automated assistance from final accountability. This structure allows artificial intelligence to support extraction, prediction, prioritization, and drafting without transferring full decision authority to the system. The evaluation of the framework should cover technical, operational, financial, and governance outcomes. Technical metrics measure predictive performance, while operational and financial metrics assess workflow effects and economic outcomes. Governance metrics assess oversight, auditability, subgroup performance, and system monitoring.

Table 3. Multilevel evaluation framework for AI-enabled RCM.

Level	Core Metrics	Example RCM Question	Minimum Expectation
Technical	AUROC, AUPRC, precision, recall, F1-score, and calibration	Does the model discriminate high-risk claims?	Temporal or external validation where feasible
Operational	Clean-claim rate, denial rate, rework, turnaround time, and staff touches	Does AI change workflow performance?	Prospective comparison or controlled rollout
Financial	Net collections, avoidable write-offs, recovered revenue, and cost-to-collect	Does AI create measurable financial value?	Risk-adjusted financial evaluation
Governance	Overrides, incidents, subgroup performance, and audit completeness	Is AI safe, fair, and accountable?	Continuous monitoring and documented escalation

4.8. Governance, Implementation, and Evidence Gaps

AI-enabled revenue cycle management processes healthcare information and financial decisions within the same workflow. The governance structure must therefore address privacy, accountability, transparency, auditability, and system monitoring throughout the implementation lifecycle. Healthcare AI governance should include clear responsibility for model outputs and human decisions. The World Health Organization identifies autonomy, safety, transparency, accountability, equity, and sustainability as relevant principles for artificial intelligence in health [16]. The FUTURE-AI guideline also defines fairness, universality, traceability, usability, robustness, and explainability as requirements for trustworthy healthcare AI [15].

Bias can arise when a model uses an outcome or proxy that does not represent the intended decision target. Historical administrative data can also contain patterns that reflect previous organizational practices. These patterns can influence predictions when the model is applied to future claims or payer-related decisions [17]. Privacy is another governance requirement because RCM data contain healthcare and financial information. The system should apply controlled access, audit logging, data minimization, secure processing, and defined retention policies. Generative artificial intelligence requires additional control because free-text inputs can contain information that is not required for the assigned task [18, 19]. Traceability supports both accountability and model evaluation. Each material AI recommendation should record the relevant source information, model version, timestamp, recommendation, user action, and final outcome when these elements are available. These records allow organizations to review decisions and monitor system behavior over time.

The implementation process can follow a phased structure. The first phase evaluates data readiness and data lineage. The second phase selects a measurable RCM use case with a defined intervention pathway. The third phase evaluates the model without changing the operational workflow. The fourth phase introduces human-reviewed recommendations and feedback mechanisms. The fifth phase evaluates controlled scaling against operational and financial outcomes. The final phase maintains monitoring, audit records, validation procedures, and retraining criteria. The reviewed evidence shows that artificial intelligence has been examined across administrative records, claims prediction, documentation, coding, denial management, and anomaly detection. However, evidence that evaluates a complete AI-enabled RCM framework across technical, operational, financial, and governance outcomes remains limited. This evidence gap supports the need for prospective evaluation in operational RCM settings.

4.9. Discussion of Evidence Maturity and Research Implications

The reviewed evidence shows different levels of maturity across AI applications in revenue cycle management. Administrative health records and claims data support predictive modeling for several operational tasks. Documentation and coding studies also report established applications for information extraction, code recommendation, and workflow support. Evidence for end-to-end financial improvement remains limited.

Current studies do not consistently evaluate technical performance, workflow outcomes, denial reduction, net revenue, and cost-to-collect within the same implementation. This limitation prevents direct conclusions about the financial impact of a complete AI-enabled RCM system.

Model performance can vary across healthcare organizations. Differences in payer composition, coding practices, contract structures, information systems, and denial processes can affect model outputs after deployment. External validation is required before a model is applied to a different operational setting. The proposed framework places artificial intelligence within the RCM workflow. The framework separates predictive output from operational action. Human staff retain responsibility for decisions that involve financial, administrative, or compliance consequences. The evaluation of AI-enabled RCM should include technical, operational, financial, and governance outcomes. Technical metrics measure model performance. Operational metrics measure workflow effects. Financial metrics measure revenue-related outcomes. Governance metrics measure oversight, auditability, and accountability. Generative artificial intelligence requires additional control within RCM workflows. Current evidence mainly covers summarization, information extraction, draft generation, and explanation tasks. Human verification should remain part of the workflow before any consequential administrative or financial action is completed.

4.10. Limitations

This study has several limitations. First, the study uses an evidence-informed integrative review rather than a registered systematic review. The review therefore does not represent an exhaustive collection of the available literature. Second, the reviewed evidence includes different data sources, methods, outcomes, and implementation settings. This heterogeneity limits direct comparison across studies. Third, part of the evidence comes from related domains, including administrative claims prediction, clinical documentation, coding, and fraud detection. These studies provide transferable evidence but do not directly evaluate a complete revenue cycle management workflow. Fourth, several implementation and financial findings originate from healthcare organizations or industry reports rather than controlled peer-reviewed evaluations. These sources provide operational information but do not replace prospective comparative evidence. Fifth, the proposed framework has not been evaluated in a live revenue cycle management environment. The study therefore does not provide prospective evidence on denial reduction, workflow outcomes, financial outcomes, or governance performance. These limitations define the scope of the findings and identify areas for future empirical evaluation.

5. Conclusion and Future Work

This study presents an evidence-informed framework for artificial intelligence in revenue cycle management. The reviewed evidence covers administrative health records, claims analytics, clinical documentation, coding, prior authorization, denial management, anomaly detection, explainable AI, and healthcare AI governance. The synthesis shows that artificial intelligence has been applied to documentation support, coding assistance, risk prediction, denial prioritization, and anomaly detection. Direct evidence for end-to-end financial improvement in revenue cycle management remains limited. Current studies do not consistently evaluate technical, operational, financial, and governance outcomes within the same implementation. The proposed framework consists of four layers: data and interoperability, predictive intelligence, workflow orchestration, and governance with human oversight. The framework separates predictive output from operational action. Human staff retain responsibility for decisions that involve financial, administrative, or compliance consequences. Future research should evaluate the framework in operational revenue cycle management settings. Studies should examine denial-risk prediction, AI-assisted coding, threshold selection, financial value, model stability, explainability, and governance controls. Future evaluations should also include different payers, healthcare facilities, clinical specialties, and reimbursement policies. Prospective studies should measure technical, operational, financial, and governance outcomes within the same implementation. Future research should also develop common benchmark datasets and consistent outcome definitions to support comparison across AI-enabled revenue cycle management studies.

Author Contributions

Conceptualization; methodology; literature synthesis; framework development; writing—original draft; writing—review and editing: Remon Gameel Ebed Saweris.

Funding

This research received no external funding.

Data Availability Statement

No new patient-level dataset was created or analyzed. The paper synthesizes published literature and publicly accessible guidance.

Institutional Review Board Statement

Not applicable. This study did not involve human participants, identifiable patient data, animals, or access to protected health information.

Acknowledgments

None.

Conflicts of Interest

The author declares no conflict of interest.

Generative AI Disclosure

Generative AI tools were used during manuscript drafting and language refinement to improve readability under human oversight. No generative AI tool was used to create, alter, or manipulate figures or other visual content in this manuscript. The author remains fully responsible for the accuracy, integrity, interpretation and final submitted version.

References

- [1] Caruana A, Bandara M, Musial K, Catchpoole D, Kennedy PJ. Machine learning for administrative health records: a systematic review of techniques and applications. *Artificial Intelligence in Medicine*. 2023;144:102642.
- [2] MacKay EJ, Stubna MD, Chivers C, Draugelis ME, Hanson WJ, Desai ND, et al. Application of machine learning approaches to administrative claims data to predict clinical outcomes in medical and surgical patient populations. *PLoS One*. 2021;16(6):e0252585.
- [3] Ravaut M, Sadeghi H, Leung KK, Volkovs M, Kornas K, Harish V, et al. Predicting adverse outcomes due to diabetes complications with machine learning using administrative health data. *NPJ Digital Medicine*. 2021;4:24.
- [4] American Hospital Association. 3 Ways AI Can Improve Revenue-Cycle Management; 2024.
- [5] U S Government Accountability Office. Science & Tech Spotlight: AI for Medical Notes and Coding. U.S. Government Accountability Office; 2026. GAO-26-109116.
- [6] Mousavi T, Shaya FT, Cooke CE. Artificial intelligence in prior authorization and coverage decisions: a systematic review of methods, evidence gaps, and future implications for patient access. *Journal of Managed Care & Specialty Pharmacy*. 2026;32(9):1135-53.
- [7] Bracken A, Reilly C, Feeley A, Sheehan E, Merghani K, Feeley I. Artificial intelligence (AI)-powered documentation systems in healthcare: a systematic review. *Journal of Medical Systems*. 2025;49(1):28.
- [8] Lee C, Britto S, Diwan K. Evaluating the impact of artificial intelligence (AI) on clinical documentation efficiency and accuracy across clinical settings: a scoping review. *Cureus*. 2024;16(11):e73994.

- [9] Zhao J, Liu H, Chen Y, Song F. Application of artificial intelligence tools and clinical documentation burden: a systematic review and meta-analysis. *BMC Medical Informatics and Decision Making*. 2025;26:29.
- [10] Mousavi Baigi SF, Sarbaz M, Darroudi A, Dahmardeh Kemmak F, Norouzi Aval R, Kimiafar K. Artificial intelligence-based automated International Classification of Diseases coding: a systematic review. *Journal of Medical Signals and Sensors*. 2025;15:22.
- [11] Li X, Zhang Y, Hou X, Wang S, Lin H. Deep learning for automatic ICD coding: review, opportunities and challenges. *Artificial Intelligence in Medicine*. 2025;168:103187.
- [12] Zhang Y, Lyu C, Chang L, Yang H, Ji B, Wei L. A systematic review of automated International Classification of Diseases coding models using the Medical Information Mart for Intensive Care dataset. *Digital Health*. 2025;11:20552076251404518.
- [13] Owolabi T. Transforming appeal decisions: machine learning triage for hospital admission denials. *JAMIA Open*. 2025;8(1):ooaf016.
- [14] du Preez A, Bhattacharya S, Beling P, Bowen E. Fraud detection in healthcare claims using machine learning: a systematic review. *Artificial Intelligence in Medicine*. 2025;160:103061.
- [15] Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ*. 2025;388:e081554.
- [16] World Health Organization. Ethics and governance of artificial intelligence for health. Geneva: World Health Organization; 2021.
- [17] Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-53.
- [18] Roberts LJ, Jayasena R, Khanna S, Arnott L, Lane P, Bain C. Challenges for implementing generative artificial intelligence (GenAI) into clinical healthcare. *Internal Medicine Journal*. 2025;55(7):1063-9.
- [19] World Health Organization. Regulatory considerations on artificial intelligence for health. Geneva: World Health Organization; 2023.
- [20] Cabitza F, Campagner A, Soares F, García de Gadiana-Romualdo L, Challa F, Sulejmani A, et al. The importance of being external: methodological insights for the external validation of machine learning models in medicine. *Computer Methods and Programs in Biomedicine*. 2021;208:106288.
- [21] Yuan H. Toward real-world deployment of machine learning for health care: external validation, continual monitoring, and randomized clinical trials. *Health Care Science*. 2024;3(5):360-4.
- [22] Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *Journal of Biomedical Informatics*. 2017;73:14-29.
- [23] Liu L, Blake V, Barman M, Gallego B, Churches T, Kennedy G, et al. Using natural language processing to extract information from clinical text in electronic medical records for populating clinical registries: a systematic review. *Journal of the American Medical Informatics Association*. 2026;33(2):484-99.
- [24] Khalid AR, Ali H, Fathima K, Kouayep SC, Kim HC. Recent advances in AI for automated ICD coding: a systematic literature review. *Journal of Medical Systems*. 2026;50(1):113.
- [25] Bedi S, Liu Y, Orr-Ewing L, Dash D, Koyejo S, Callahan A, et al. Testing and evaluation of health care applications of large language models: a systematic review. *JAMA*. 2025;333(4):319-28.
- [26] Wang L, Wan Z, Ni C, Song Q, Li Y, Clayton E, et al. Applications and concerns of ChatGPT and other conversational large language models in health care: systematic review. *Journal of Medical Internet Research*. 2024;26:e22769.
- [27] Jung J, Lee H, Jung H, Kim H. Essential properties and explanation effectiveness of explainable artificial intelligence in healthcare: a systematic review. *Heliyon*. 2023;9(5):e16110.

- [28] Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. *JMIR AI*. 2024;3:e53207.
- [29] Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digital Medicine*. 2018;1:18.
- [30] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *New England Journal of Medicine*. 2019;380(14):1347-58.
- [31] Ben-Israel D, Jacobs WB, Casha S, Lang S, Ryu WHA, de Lotbiniere-Bassett M, et al. The impact of machine learning on patient care: a systematic review. *Artificial Intelligence in Medicine*. 2020;103:101785.