

Salmona: A Multi-Domain Architecture for Traceable Medical Billing Review in Colombian Healthcare

Katherine Monsalve

katherine.monsalve@arkangel.ai

Arkangel AI

Natalia Castaño-Villegas

Arkangel AI

Fiorella Ramirez

Arkangel AI

Elmer Raúl Escandón

Arkangel AI

Jose Zea

Arkangel AI

Laura Velásquez

Arkangel AI

Research Article

Keywords: Insurance Claim Review, Artificial Intelligence, Insurance, Health, Medical Necessity, Clinical Decision Support Systems, Colombia, Medical Informatics, Expert Systems

Posted Date: June 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9819774/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: Competing interest reported. The authors are employees of Arkangel AI, the company that developed the Salmona architecture described in this manuscript. The authors declare no other financial or non-financial competing interests.

Abstract

Background

Medical billing audit in Colombia requires the simultaneous verification of administrative validity, clinical pertinence, and financial consistency of each claim, and may extend for more than three months per batch of submitted claims.

Published artificial intelligence systems for this process focus on financial fraud detection through black-box classifiers, without covering the three audit domains simultaneously or describing concrete mechanisms of human governance.

Methods

A descriptive methodological study was conducted to document the design, architecture, and decision rationale of Salmona, a rule-grounded medical billing audit pipeline developed by Arkangel AI for healthcare reimbursement workflows in the Colombian health system. The description is based on the system's technical documentation and the applicable Colombian regulatory framework. No patient data were collected and no empirical validation was conducted in this study.

Results

Salmona decomposes audit into three independent specialized audit layers: an administrative layer (27 rules, DAMA-UK instrument), a clinical pertinence layer (29 rules, PERT-CLIN instrument), and a financial and contractual layer (42 rules, FIN-CTR instrument). Each rule produces a structured result with three possible values: pass when there is evidence of compliance, fail when there is evidence of a violation, and n/a when the necessary information is absent. This innocent-until-proven-otherwise principle prevents the absence of a document from automatically becoming a billing objection. A canonical consolidation layer reconciles findings per invoice item, assigns objection causals according to Technical Annex 6 of Resolution 3047 of 2008, and calculates objected amounts under a financial integrity invariant. The human auditor reviews the consolidated result, may iteratively request corrections via comments, and initiates report delivery to the provider through an explicit action that is independent of the review cycle.

Conclusions

Salmona constitutes, to our knowledge, an early methodological contribution to the description of multi-domain medical billing review systems for the Colombian health system, with per-rule traceability and explicit human governance.

Although initially oriented toward Colombian health insurers, the described architecture may also support provider-side pre-submission validation and structured response to reimbursement objections. The described architecture provides the basis for empirical validation with real claims in future studies.

Background

Errors and fraud in healthcare service billing represent between 3% and 15% of total global health expenditure, generating losses that compromise the financial sustainability of health systems and reduce the availability of resources for patient care.[1] More than 80% of medical claims are estimated to contain some type of error, whether due to inconsistencies in the coding of diagnoses and procedures, absence of supporting documentation, or discrepancies between amounts billed and existing contracts.[1, 2] In clinical coding specifically, manual error rates can reach 20%, with an estimated economic impact in the billions of dollars annually even in systems with greater automation than those in Latin America.[3]

In Latin America, the complexity of health systems, characterized by the coexistence of multiple insurance regimes, heterogeneous fee schedules, and continuously updated regulatory frameworks, amplifies the impact of billing errors and makes their systematic detection more difficult.[6, 7] Payers in the region face growing volumes of claims accompanied by unstructured clinical documentation, in contexts where the installed capacity for human audit is limited and traceability mechanisms are incipient.[6, 7]

In Colombia, this phenomenon is expressed with particular intensity: billing objections (glosas), that is, formal objections to medical claims submitted by providers, can represent a substantial proportion of the billing of a healthcare institution, and the accumulated debt of health-promoting entities (EPS) to providers has been documented in the multi-trillion-peso range.[5, 6, 7]

The Colombian regulatory framework establishes particularly complex conditions for this process. Resolution 2275 of 2023, amended by Resolution 1884 of 2024, establishes mandatory integration between the electronic sales invoice and the Individual Health Service Provision Registry (RIPS), advancing toward the digitization of the claims submission process.[8] Nevertheless, the audit process following submission remains predominantly manual and can extend for more than three months per batch of claims (i.e., a group of claims submitted together in a single billing cycle), deteriorating the cash flow of healthcare institutions and generating operational inefficiencies in health-promoting entities.[9] In response to this challenge, the national government launched in 2026 the Intelligent Audit System of the Administrator of Resources of the General System of Social Security in Health (SIA-ADRES), with the objective of reducing processing times to less than two weeks; however, it initially operates on special-event claims and no academic publication describing its architecture in detail exists to date.[9]

Artificial intelligence has emerged as an alternative for automating and improving the efficiency of healthcare audit and billing processes.[2] The most documented applications include fraud detection through machine learning models, with reported improvements in claim anomaly detection compared to traditional methods[10]; natural language processing for extraction and classification of claims information at scale[21]; and multi-agent artificial intelligence systems, in which multiple specialized agents collaborate to resolve complex tasks in clinical and administrative environments, with reported improvements in efficiency and diagnostic accuracy.[4, 20]

The Arkangel AI team has previously described a multi-agent architecture for automated clinical coding in Colombia[14], whose design preceded and informed the development of the system described in this study.

However, medical billing audit in the context of healthcare payers and providers is a multi-domain workflow that simultaneously requires verifying the administrative completeness of the record, evaluating the clinical pertinence of services rendered, validating amounts billed against the contract, and justifiably documenting each objection raised to the provider. The majority of published artificial intelligence approaches address anomaly or financial fraud detection as the primary objective, without covering the other domains of the audit process.[10, 11] These systems frequently operate as black-box classifiers: they produce a score without exposing the applied rules or citing the evidence supporting each finding, which hinders review by the human auditor and regulatory traceability.[17, 19] Regarding human oversight, although the literature recognizes it as a necessary condition for ensuring safety and legitimacy in artificial intelligence systems in health, concrete oversight mechanisms are rarely described with precision: the presence of a human in the process is mentioned without specifying what authority they hold, at what moment they intervene, or how their decisions are recorded.[12, 13]

In the Colombian health system, no published description of a structured multi-domain audit system exists to date that simultaneously addresses the administrative, clinical, and financial domains of medical billing audit, with traceable rules, structured outputs per invoice item, and an explicit mechanism for auditor-supervised correction. Therefore, the objective of the present study was to describe the design, architecture, and decision rationale of Salmona, a medical

billing audit system developed for healthcare reimbursement workflows in Colombia, as an early methodological contribution, to our knowledge, in this line of research in the context of the Colombian health system.

Methods

2.1 Study design

A descriptive methodological study was conducted to document the design, architecture, and decision rationale of Salmona. This type of study does not involve patient data collection, does not require ethics committee approval, and does not include empirical validation metrics. The contribution of the present study is the architectural and methodological description of the system: the decomposition of the audit problem into independent domains, the rule evaluation logic, and the human oversight mechanism. As no specific reporting guidelines exist for studies describing medical audit architectures of this type, the present manuscript was informed by the principles of TRIPOD-AI and CONSORT-AI, as well as existing frameworks for the responsible implementation and governance of artificial intelligence systems in health.[24]

2.2 Development context

Salmona was developed by the technical and scientific team of Arkangel AI, a Colombian artificial intelligence company in health. Development was informed by co-design sessions with audit teams from Colombian health-promoting entities, whose input guided the definition of review domains and the rules applied in each layer. At the time of writing, Salmona is in operational validation in real medical billing audit and claims management scenarios within the Colombian health system, with capacity to support both payer-side and provider-side workflows.

2.3 Regulatory framework informing the design

Resolution 3047 of 2008 of the Ministry of Social Protection establishes the Single Manual of Objections, Returns, and Responses, which defines seven billing objection causals (Technical Annex 6) under which a payer may formally object to a claim.[15] These causals structure the system's consolidation layer and the mapping of findings to specific objections. Resolution 2275 of 2023, amended by Resolution 1884 of 2024, establishes mandatory integration between the electronic invoice and the Individual Health Service Provision Registry.[8] The clinical practice guidelines of the Ministry of Health and Social Protection inform the clinical pertinence layer when available for the main diagnosis of the claim.

2.4 Audit rule derivation process

The three audit layers comprise a total of 98 rules distributed across three independent instruments. The 27 rules of the administrative layer (DAMA-UK instrument) were constructed by adapting the seven data quality dimensions of the standard DAMA International framework (accuracy, completeness, consistency, validity, uniqueness, timeliness, and traceability) to the medical claims submission process under the Colombian regulatory framework: Law 100 of 1993, Resolution 3047 of 2008, Decree 4747 of 2007, and Resolution 1536 of 2015.[15, 16] Each dimension was operationalized into rules verifiable against the claim documents and the administrative records of the system, including consultation of the Unique Affiliates Database (BDUA).

The 29 rules of the clinical layer (PERT-CLIN instrument) operationalize the concept of clinical pertinence using the Clinical Practice Guidelines of the Ministry of Health and Social Protection as the decision reference. When the ICD-10 code of the case does not have a guideline available in the reference corpus, the guideline-dependent rules adopt an n/a result and the case is escalated to the human auditor.

The 42 rules of the financial layer (FIN-CTR instrument) cover tariff and contractual audit differentiated by benefit plan (for example, plans with different coverage levels) and include specific anti-fraud controls: duplicate billing across multiple provider institutions, double billing across regimes (SOAT/EPS/ARL), upcoding (billing higher-value codes than the procedure actually performed), and unbundling (improper disaggregation of procedures that should be billed as a package). Representative rule examples per layer are presented in Table 4. The rules of each layer can be adjusted according to the specific contract and coverage of each client, without modifying the central logic of the pipeline.

The rule set was defined by the Arkangel AI technical team and was not submitted to external review by medical auditors or health-promoting entity experts outside the development team. This limitation is discussed in the corresponding section.

2.5 Technological infrastructure

Each Salmona audit layer is executed by an artificial intelligence agent that uses OpenAI Codex (OpenAI, San Francisco, CA; 2025), accessed via the OpenAI API. Codex operates in reasoning mode, in which the model applies internal chain-of-thought before emitting each response; this mode does not use a temperature parameter and is configured to produce highly reproducible structured outputs for a given input. The agents generate their results through structured model calls with JSON schema enforced in the system prompt, which directly produces the output fields in the specified format. The system operates under a multi-agent framework with tool use capability, allowing each agent to consult documents, execute searches in reference databases, and emit structured results as part of a coordinated workflow.

The system receives claim documents in four formats: XML (for structured data such as RIPS and invoice data), CSV (for tabular records), PDF (for scanned or digitally generated clinical documents), and Word or other structured text formats (for clinical records and narrative supports). The content of each format is processed using structured reading to extract fields and normalized before analysis by the audit agents.

2.6 Separable knowledge model

Salmona separates fixed audit logic from client-specific reference knowledge. Fixed logic comprises the rules of the three layers, severity and weight criteria, result semantics, and billing objection causal mapping. Configurable knowledge includes clinical practice guidelines, contractual fee schedule manuals, and benefit plan coverage rules. The clinical layer can continue operating without available guidelines (dependent rules adopt n/a). The financial layer requires the fee schedule files; their absence halts the process with a descriptive error.

This separation between fixed logic and configurable knowledge is the property that makes the Salmona architecture adaptable to other health systems: in a different regulatory context, adaptation would begin with replacing the regulatory and clinical reference files with their local equivalents, including national clinical guidelines, system-specific fee schedule manuals, and applicable billing regulations, without modifying the central pipeline logic. In a Latin American health system with comparable structures, for instance, the fee schedules and clinical guidelines of that system would replace those of Colombia without altering the audit rules.

Prior to operational deployment, the pipeline was subjected to functional testing using 55 synthetic billing cases generated deterministically (seed = 42), covering ten clinical scenarios across four fictional provider institutions. The test set yielded the following expected output distribution: 35 compliant cases (APTA, no objection findings), 10 cases with structured objection findings (NO_APTA), 5 cases resulting in administrative return to provider triggered by structural defects at the administrative gate, and 5 cases escalated to mandatory human review due to absence of an applicable clinical practice guideline. Full details are provided in Supplementary Material 1.

Results

3.1 Overall pipeline overview

Salmona receives a claim submitted by a healthcare provider institution, processes it through three independent audit layers, and produces a canonical structured result per invoice item, reviewed by a human auditor before issuing any communication to the provider. The pipeline operates in sequence: (1) case intake and structuring; (2) administrative audit; (3) clinical audit; (4) financial audit; (5) canonical consolidation; (6) review and correction by the human auditor; (7) report delivery to the provider as an independent action. See Fig. 1.

3.2 Case intake and context

Each case is initiated with a structured audit context containing: case identifier, submission date, invoice number, provider and payer identification, list of attached documents, and workflow status flags. This context is the single source of truth for all pipeline layers. Attached documents are processed according to their format: RIPS XML files are processed using structured reading to extract structured fields; PDF, Word, and CSV documents are processed to extract their textual content. In this way, each layer accesses the relevant content in a readable format, regardless of the original document format.

3.3 Independent audit layers

Salmona decomposes audit into three independent specialized audit layers (Table 1). The clinical layer does not access the administrative result, and the financial layer does not access the clinical result. This separation prevents cross-domain contamination of reasoning. Each layer emits a rule-level structured result including: identifier, name, severity, weight, cited evidence, observations, and result (pass, fail, or n/a). Representative examples of rules from each layer are presented in Table 4.

3.3.1 Administrative layer (DAMA-UK)

The administrative layer evaluates whether the claim record meets the formal requirements to be audited. Its 27 rules cover: patient identity consistency, RIPS structure, presence of authorization, completeness of supporting documents, timeliness of submission, and record traceability. If this layer produces a return indicator, Salmona generates the final result and halts the process.

3.3.2 Clinical layer (PERT-CLIN)

The clinical layer evaluates whether billed procedures are clinically pertinent, documented in the clinical record, and consistent with applicable clinical practice guidelines. Its 29 rules cover: clinical indication, justification in the clinical record, medication pertinence, diagnostic aids, and hospitalization criteria. If no applicable CPG is available, dependent rules adopt n/a with an explanatory observation and the case is escalated to the human auditor.

3.3.3 Financial layer (FIN-CTR)

The financial layer evaluates whether billed amounts correspond to the contract, the applicable fee schedule, the benefit plan coverage, and anti-fraud controls. Its 42 rules cover: verification of rates by CUPS code, supporting fee schedule manuals, coverage and plan exclusions, copayments, moderating fees, and detection of duplicate charges. The absence of financial reference files is a blocking condition.

Table 1
Salmona audit layers

Layer	Instrument	Rules	Central question	Role in the pipeline
Administrative	DAMA-UK	27	Is the record complete, valid, timely, and traceable?	Verification checklist and return gate
Clinical	PERT-CLIN	29	Are the billed services clinically necessary, documented, and consistent with CPGs?	Verification checklist
Financial	FIN-CTR	42	Do the billed amounts correspond to the contract, fee schedule, coverage, and anti-fraud controls?	Verification checklist

3.4 Innocent-until-proven-otherwise logic

Salmona explicitly distinguishes three per-rule results: (i) pass, when there is positive evidence of compliance; (ii) fail, when there is positive evidence of a violation; and (iii) n/a, when the necessary information is absent or the rule does not apply. This distinction prevents the absence of a document from automatically becoming a billing objection. If the system does not find a required support document, it records an observation for the human auditor without penalizing the provider.[15] This distinction has direct practical implications: in the Colombian context, the automatic attribution of a billing objection due to documentary absence is one of the documented causes of prolonged disputes between health insurers and providers. Salmona implements this principle to separate the objective verification of inconsistencies from automatic penalization for information gaps.

3.5 Canonical consolidation

Following the three layers, consolidation reconciles findings into a canonical result centered on the invoice item. The process follows four steps: (1) separation of findings (fail) and observations (n/a due to missing information); (2) grouping by item (CUPS code and service date); (3) prioritization by severity and objected amount; and (4) causal assignment according to Technical Annex 6 of Resolution 3047 of 2008.[15] Step (4) applies the predefined hierarchy of seven causals (Table 2): $4 > 2 > 5 > 1 > 3 > 6 > 7$. When multiple rules from different domains affect the same invoice item, the system assigns a single primary causal according to this predefined hierarchy. The hierarchy prioritizes inconsistencies of greater financial impact, such as duplicate billing (causal 4) and lack of clinical pertinence (causal 2), over secondary discrepancies such as tariff deviations or documentary defects, preventing fragmented or contradictory objections regarding the same billed service.

Table 2
Billing objection causal mapping per Technical Annex 6, Resolution 3047 of 2008

Causal	Meaning	Typical triggers in Salmona	Priority
1	No contractual coverage	Plan exclusions, incorrect modality, uncovered service	4
2	No clinical pertinence	CPG deviations, unjustified procedures or medications	2
3	Incomplete documentation	Missing authorization, clinical support, operative note, or annex	5
4	Duplicate or conflicting charge	Double billing, overlaps, improper disaggregation	1
5	Incorrect tariff	Manual, UVB/UVR, base rate, settlement, copayments	3
6	Coverage exhausted	Caps, waiting periods, exhausted benefits	6
7	Submission defect	Dates, consecutive numbering, general traceability	7

3.6 Per-item output model (Hallazgoltem)

The audit unit of Salmona is each billable item identified by CUPS code and service date. The canonical per-item result is represented in the Hallazgoltem object (Table 3). The global result includes: total billed amount, total approved amount, total objected amount, final concept (APPROVED or NOT_APPROVED), required action, and executive summary.

Table 3
Fields of the per-item output model (Hallazgoltem)

Field	Description
cups_code	Code of the billed procedure or service
description	Human-readable description of the service
billed_amount	Amount charged for the item
finding	Item status: compliant or objection
layer	Finding source layer: administrative, clinical, or financial
applied_rule	Determining or most severe rule associated with the item
severity	Finding severity: critical, major, or minor
objected_amount	Objected amount for the item
amount_to_recognize	Amount to be recognized to the provider after the objection
suggested_objection	Structured causal, text, amount, and currency of the proposed objection
required_evidence	Additional evidence requested from the provider or auditor

3.7 Two-bucket financial integrity invariant

The system applies a financial integrity constraint requiring that the sum of the approved amount and the objected amount equals the total billed amount. This invariant is calculated at the item level, preventing double-counting when multiple rules flag the same service. After each modification in the human review cycle, the system recalculates and verifies that the invariant is maintained.

3.8 Auditor-supervised correction cycle

Following consolidation, the canonical result enters human review. The system classifies each case into one of two states: Ready, when the system determines through its internal reasoning process that the result is consistent and ready for review; or Requires Review, when the system identifies ambiguities or inconsistencies that must be resolved before presenting the result. In both states, the human auditor reviews the result in their work interface and makes the final decision.

If the auditor agrees with the findings, they accept the result without further action. If the auditor identifies an error or disagrees with a finding, they request a revision via a comment in the interface. The system processes the comment and applies the corresponding modifications to the canonical result: updating values, tables, and monetary aggregates as indicated. The financial integrity invariant is automatically re-verified after each modification. Each change is recorded in an audit log with actor, applied operations, and timestamp, ensuring the traceability of all human auditor decisions.[12, 13]

After the modification, the human auditor reviews the updated result. This cycle repeats until the result is accepted without pending corrections. Report delivery to the provider constitutes a completely independent workflow from the review cycle: the user initiates it explicitly through an action in the interface, the system groups the objections to be sent,

generates the report in spreadsheet format, and sends it by email to the provider. The system does not send communications autonomously under any circumstances. See Fig. 2.

Table 4
Representative rule examples per audit layer

ID	Layer	Name	Severity	Required evidence	Pass	Fail	N/a
A05	Administrative	Active rights	Critical	BDUA query	Affiliate active at time of care	Affiliate inactive or not found	BDUA unavailable
A08	Administrative	Valid authorization	Critical	Authorization support	Valid and current authorization number	Invalid number or expired authorization	Emergency case (exempt)
M04	Clinical	Current CPG applied	Critical	Clinical record + MinSalud CPG	Care aligned with applicable CPG	Deviation without clinical justification	ICD-10 without available CPG (escalated to human)
M14	Clinical	Medication pertinent to diagnosis	Critical	Clinical record, medical order, CPG	Medication corresponds to indication	Medication does not correspond to diagnosis	Indication not documented
F07	Financial	Agreed fee schedule	Critical	Contract and fee schedule annex	Contract fee schedule in force is used	ISS 2001 used when specific contract exists	Contract not loaded or without annex
F13	Financial	Correct base rate	Critical	Fee schedule and billing tool	Billed rate matches fee schedule	Billed rate differs from fee schedule	CUPS code not found in fee schedule

Discussion

The description of Salmona as a medical billing review system for the Colombian health system represents an early methodological contribution to the description of architectures for medical billing audit. The system simultaneously addresses the three domains of medical billing audit, administrative, clinical, and financial, with traceable rules, structured outputs per invoice item, and concrete human governance mechanisms, responding to the need of audit teams in Colombian health-promoting entities for methodologically verifiable tools in a system whose financial unsustainability has been documented.[6, 7]

4.1 Distinctive characteristics compared to published systems

Unlike the majority of published artificial intelligence systems for medical billing audit, which address financial fraud detection as a single problem through machine learning classifiers[10], Salmona treats audit as a multi-domain problem. Clinical discrepancies (causal 2) and administrative discrepancies (causals 3 and 7) are equally frequent and have distinct regulatory implications that a black-box classifier cannot express with the granularity required by Technical Annex 6 of Resolution 3047 of 2008.[15] Additionally, while fraud detection systems produce a risk score to prioritize which claims to review, the objective of Salmona is not to automate invoice rejection, but to structure the clinical, administrative, and financial consistency review before inconsistencies become formal objections or disputes.

Multi-agent systems have begun to be explored in clinical and administrative environments as an evolution of single-agent approaches.[4, 20] The Salmona architecture extends this paradigm to medical billing audit, for which no equivalent published descriptions oriented to healthcare billing review workflows were found.

The published system closest in design to Salmona is IBM Research Europe's Clais, which extracts rules from health policy documents and allows the auditor to edit them.[11] However, Clais does not report decomposition into independent domains or logic for distinguishing confirmed violations from missing information. In Colombia, the SIA-ADRES operates on special-event claims on the public payer side without an academic publication of its architecture.[9] Both systems are complementary.

4.2 Design decisions and their rationale

The explicit rule-based approach aligns with the clinical decision support systems literature, which indicates that these systems are the most mature application of artificial intelligence in clinical environments due to their transparency, interpretability, and adaptability, and that domain separation reduces the risk of cross-contamination of reasoning in multi-criteria systems.[17] Layer independence is also consistent with the multi-agent systems in health literature: specialized agents reduce cross-domain bias and facilitate human review of each finding independently.[4, 20]

The innocent-until-proven-otherwise logic directly responds to the explainability literature in artificial intelligence in health: clinical systems must precisely distinguish what they can evaluate and what they cannot, and information gaps must be visible to the human reviewer.[19]

The auditor-supervised correction cycle implements the requirements that the literature identifies for effective human governance in artificial intelligence systems in health: decision traceability, change auditing, and conservative behavior under ambiguity.[12, 13, 18, 24] The auditor controls two explicit moments — iterative review, in which they can accept the result or request corrections via comments, and delivery as an independent action — ensuring that no communication reaches the provider without their conscious approval.

4.3 Strengths and limitations

Among the strengths of the present manuscript are the detailed description of a novel architecture for the Colombian health system context, the explicit alignment of the design with the applicable regulatory framework, the separation between fixed and configurable knowledge that facilitates scalability without rewriting the central system logic, and the concrete specification of the human oversight mechanism with two differentiated moments (review and delivery) that responds to limitations widely identified in the literature.[12, 13, 18]

The main limitation is the absence of empirical validation with real claims. The preliminary tests employed synthetic billing scenarios constructed by the technical team, which introduces a risk of design bias: the evaluated scenarios reflect the team's understanding of possible cases and not the actual distribution of claims in a health-promoting entity. The literature warns that excessive reliance on synthetic tests can propagate biases and compromise the generalizability of artificial intelligence systems in health.[22] This limitation was partially mitigated through the controlled design of representative scenarios covering the main billing objection causals. Prospective validation with real claims from Colombian health-promoting entities is planned as the next phase.

A second limitation is the dependence on the quality of the content extraction process from claim documents. Clinical documents in Colombia present variability in format, legibility, and degree of digitization, which may affect information extraction and increase the proportion of rules with an n/a result.[23] The supervised correction cycle acts as a safeguard, given that the auditor verifies the result before any communication to the provider.

A third limitation is that the rule set was defined by the Arkangel AI technical team and was not submitted to external review by medical auditors or health-promoting entity experts outside the development team. The complexity of medical billing validation in Colombia and the need for systematic tools for managing billing objections have been previously documented.[25] Manual coding errors by human professionals can reach 20%.[3] As partial mitigation, the rules were oriented by co-design sessions with audit teams from Colombian health-promoting entities during development. Validation of the rule set by independent auditors from Colombian health-promoting entities is a necessary step before large-scale implementation.

Additionally, there is a risk of regulatory obsolescence. The rules of the three audit layers were derived from the Colombian regulatory framework in force at the time of system development. Medical billing and health audit regulations are subject to continuous updating by the Ministry of Health and Social Protection and other regulatory authorities. Any modification to the Single Objections Manual (Resolution 3047 of 2008), RIPS standards (Resolution 2275 of 2023), or Clinical Practice Guidelines may render specific rules outdated, with the corresponding risk of unjustified objections or omission of legitimate findings. This risk requires a continuous process of review and updating of the system rule set.

Likewise, there is a risk of extraction error propagation. The Salmona agents use large language models to extract information from claim documents before evaluating the rules. Language models can produce incorrect, incomplete, or hallucinated extractions, particularly in unstructured clinical documents, documents with ambiguous terminology, or low typographic quality. Given that the rules of the three layers are evaluated on the extracted information, an extraction error can propagate to multiple rule results without being automatically detected by the system. The actual proportion and impact of this type of error under operational conditions are unknown until validation is conducted with genuine claims and independent reference auditors. The supervised correction cycle partially reduces this risk by subjecting each result to human review before any communication to the provider.

Finally, it is worth noting that while the empirical validation with real claims has not yet been completed, the team is in the process of establishing collaborations with Colombian health-promoting entities for a prospective evaluation that will constitute the subject of a subsequent study.

4.4 Implications

These findings suggest that artificial intelligence implementation strategies for medical billing audit that rely exclusively on financial fraud detection will have limited scope in the Colombian context, given that the three audit domains generate independent objection causals under the framework of Resolution 3047 of 2008 that require specialized reasoning.[15] The described architecture was designed with the objective of reducing inter-auditor variability by standardizing domain-level evaluation rules, and could contribute to improving the traceability of objection decisions and reducing processing time per claim; however, none of these effects have been empirically measured to date. Responsible implementation of artificial intelligence systems in health environments requires a structured evaluation process that includes technical validation, performance monitoring in real-world context, and mechanisms for continuous system maintenance.[24]

It is important to distinguish between two dimensions of Salmona: the current technological implementation and the underlying base architecture. The current implementation is Colombia-specific: the administrative rules reference Colombian regulations (Resolution 3047 of 2008, Decree 4747 of 2007, Law 100 of 1993), the clinical layer uses the Clinical Practice Guidelines of the Colombian Ministry of Health and Social Protection, and the financial layer operates with the fee schedules and benefit plans of the Colombian health system. However, the base architecture of the system, comprising the three independent audit layers, the pass/fail/n/a rule logic, the separable knowledge model, and the human oversight cycle, is independent of the specific regulatory context and can be replicated in other health systems.

In a health system with comparable payer-provider structures, adaptation would begin with replacing the regulatory reference files, clinical guidelines, and fee schedules with their local equivalents, keeping the central pipeline logic intact.

The system was designed to support both payer-side and provider-side workflows within the Colombian health system. For payers, it structures the audit of submitted claims, objection generation, and administrative review. For providers, the same rule framework can be applied to pre-submission self-audit and to the preparation of structured responses to received objections. In both cases, a shared rule set enables evaluation from either side of the billing relationship, with consistent criteria and full traceability of findings.

Prospective studies are required to evaluate the performance of Salmona against real claims from Colombian health-promoting entities, using the consensual judgment of expert auditors as a reference standard, and to measure the effects of its implementation on audit times, objection rates by causal, concordance between the system and the human auditor, and financial impact for the payer and provider. Additionally, future studies should explore the adaptability of the architecture to other health systems in the Latin American region, where the health insurance model with medical billing audit presents structural characteristics comparable to the Colombian system.

Conclusions

Salmona decomposes medical billing audit into three independent specialized layers covering the administrative (27 rules), clinical (29 rules), and financial (42 rules) domains, and produces traceable per-invoice-item results with rule-based assignment of objection causals according to the Colombian regulatory framework. The distinctive architectural characteristic compared to published systems is the combination of simultaneous multi-domain audit, a three-valued rule logic (pass/fail/n/a) that distinguishes confirmed violations from information gaps, and a human oversight cycle with two explicit moments, iterative review and delivery as an independent action, that ensures no communication reaches the provider without the conscious approval of the responsible auditor. While the current implementation is oriented toward the Colombian health system, the base architecture is replicable in other healthcare contexts by replacing the regulatory and clinical reference files with their local equivalents. Prospective studies are required to evaluate the performance of Salmona against real claims, with expert auditors' judgment as the reference standard.

Declarations

Ethics approval and consent to participate

Not applicable. This study describes a system architecture and does not involve human participants, patient data, or clinical interventions.

Consent for publication

Not applicable.

Competing Interests

The authors are employees of Arkangel AI, the company that developed the Salmona architecture described in this manuscript. The authors declare no other financial or non-financial competing interests.

Funding

No external funding was received for this study.

Author Contribution

KM: Conceptualization, Methodology, Writing – original draft, Writing – review and editing. NCV: System design, Software, Writing – review and editing. FR: System design, Software, Methodology. RE: System design, Software, Validation. JZ: Project administration, Supervision. LV: Writing – review and editing.

Acknowledgements

The authors thank the EPS audit teams who participated in co-design sessions that informed the rule set definition.

Data Availability

The datasets used in this study consist of test examples created to illustrate and describe the Salmona architecture. These data do not correspond to real patient or billing records and were not used to validate the system. The example datasets are available from the corresponding author on reasonable request.

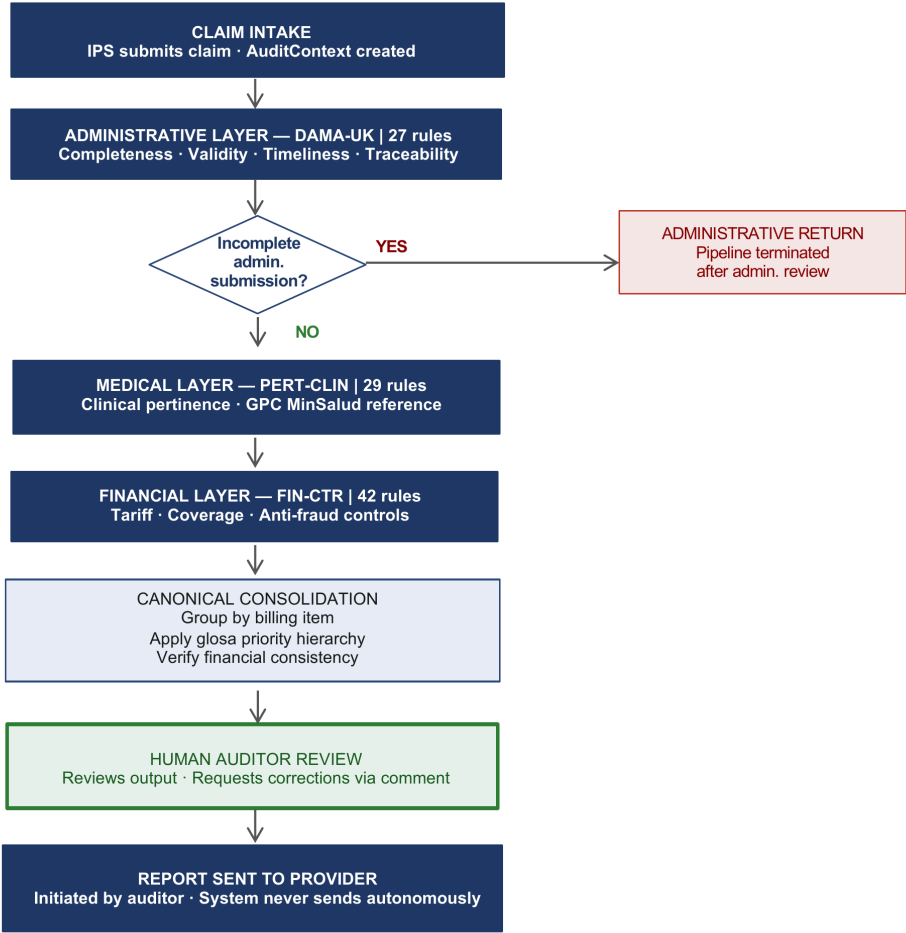
References

1. Vafaei Najar A, Alizamani L, Zarqi M, Hooshmand E. A global scoping review on the patterns of medical fraud and abuse: integrating data-driven detection, prevention, and legal responses. *Arch Public Health*. 2025;83(1):43. 10.1186/s13690-025-01512-8.
2. Nasser LK. The evolution of automated medical billing with artificial intelligence. *Cureus*. 2025;17(11):e77840.
3. Gershon A, Soffer S, Nadkarni GN, Klang E. Automatic ICD coding using LLMs: a systematic review [Preprint]. *medRxiv*. 2025. 10.1101/2025.07.30.25330916.
4. Srinivasu PN, Kumari GLA, Ahmed S, Alhumam A. Exploring agentic AI in healthcare: a study on its working mechanism. *Front Med*. 2026;12:1753443.
5. Ministerio de Salud y Protección Social de Colombia. La deuda acumulada de las EPS que asciende a \$32 billones de pesos [Internet]. Bogotá: Minsalud; 2025 [cited 2026 May 24]. Available from: <https://www.minsalud.gov.co/CC/Noticias/2025/Paginas/deuda-acumulada-de-EPS-asciende-a-32-billones-de-pesos.aspx>
6. Espinosa O, Rodríguez J, Arias M, Pinzón D, Valdez E. Analysis of the financial situation of the maximum budgets of the Health-Promoting Entities in Colombia. *World Med Health Policy*. 2024;16(4):745–58. 10.1002/wmh3.635.
7. Espinosa O, Rodríguez-Lesmes P, Prada S, Bolívar M, Mejía-Becerra JD, Vecino-Ortíz AI. Health insurance premium in Colombia for 2025: a strictly political-ideological decision without technical-scientific arguments? *Health Econ Rev*. 2025;15(1):31. 10.1186/s13561-025-00620-z.
8. Colombia. Ministerio de Salud y Protección Social. Resolución 2275 de 2023. Por la cual se modifican los estándares del sistema de información para la protección social. Bogotá; 2023. Modificada por Resolución 1884 de 2024.
9. ADRES. ADRES lanza el Sistema Inteligente de Auditoría soportado por IA [Internet], Bogotá ADRES. 2026 [cited 2026 May 24]. Available from: <https://www.adres.gov.co/sala-de-prensa/noticias/adres-lanza-el-sistema-inteligente-de-auditoria-soportado>
10. du Preez A, Bhattacharya S, Beling P, Bowen E. Fraud detection in healthcare claims using machine learning. *Artif Intell Med*. 2024;160:103061.

11. Sbodio ML, López V, Hoang TL, et al. Collaborative AI system for healthcare claims compliance. *Sci Rep*. 2024;14:12089.
12. Griffen Z, Owens K. From human in the loop to participatory governance for AI in healthcare. *Am J Bioeth*. 2024;24(9):81–3.
13. Olawade DB, Plabon SB, Ojo A, Ogunbona MA, Makanjuola BD, Olasilola OR. Human in the loop artificial intelligence in healthcare: applications, outcomes, and implementation challenges. *Int J Med Inf*. 2026;213:106362. 10.1016/j.ijmedinf.2026.106362.
14. Castaño Villegas N, Escandon R, Monsalve K, Zea J, Velasquez L. A recursive learning architecture for zero-shot automated clinical coding: a methodological study [Preprint]. *JMIR Preprints*. 2026. 10.2196/preprints.98279.
15. Colombia. Resolución 3047 de 2008. Manual Único de Glosas, Devoluciones y Respuestas. Bogotá; 2008.
16. DAMA International. DAMA-DMBOK: Data Management Body of Knowledge. 2nd ed. Technics; 2017.
17. Alnattah A, Jajroudi M, Fadafen SAN, Manzari MN, Eslami S. Artificial intelligence in clinical decision-making: a scoping review of rule-based systems and their applications in medicine. *Cureus*. 2025;17(8):e91333. 10.7759/cureus.91333.
18. Gárate FJ, Chausa P, Moreno D, López Luque J, Díaz-Brito V, Gómez EJ. A governance and evaluation framework for deterministic, rule-based clinical decision support in empiric antibiotic prescribing [Preprint]. *arXiv*. 2026. arXiv:2603.10027.
19. Zhang K, Wang D, Lin F, Xie J, Zhou W. A comprehensive review of explainable artificial intelligence in healthcare: methods, evaluation, and clinical integration. *iScience*. 2026;29(3):115026. 10.1016/j.isci.2026.115026.
20. Borkowski AA, Ben-Ari A. Multiagent AI systems in health care: envisioning next-generation intelligence. *Fed Pract*. 2025;42(5):188–94. 10.12788/fp.0589.
21. 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. *J Am Med Inf Assoc*. 2026;33(2):484–99. 10.1093/jamia/ocaf176.
22. Koul A, Duran D, Hernandez-Boussard T. Synthetic data, synthetic trust: navigating data challenges in the digital revolution. *Lancet Digit Health*. 2025;7(11). 10.1016/j.landig.2025.100924.
23. Ministerio de Salud y Protección Social de Colombia. ABC de interoperabilidad de datos de la historia clínica en Colombia [Internet]. Bogotá: Minsalud; 2023 [cited 2026 May 24]. Available from: <https://www.minsalud.gov.co/ihc/Documentos%20compartidos/ABC-IHC.pdf>
24. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO. 2021. Licence: CC BY-NC-SA 3.0 IGO. Available from: <https://www.who.int/publications/i/item/9789240029200>
25. Chavarría T. Desarrollo e implementación de una herramienta para la validación de cuentas médicas y gestión de glosas en el sector salud en Colombia. *Med UPB*. 2018;37(2):131–41. 10.18566/medupb.v37n2.a07.

Figures

Figure 1. Salmona Pipeline — General Flow



Abbreviations: IPS = healthcare provider institution; EPS = health insurer; GPC = clinical practice guidelines; CUPS = Colombian health procedure codes.

Figure 1

General flow of the Salmona pipeline. The system receives the claim submitted by the provider, processes it through three independent audit layers (administrative, clinical, and financial) in sequence, consolidates the findings into a canonical result per invoice item, and submits the result for review and correction by the human auditor. Report delivery to the provider is an independent action explicitly initiated by the user.

Figure 2. Human-in-the-Loop Review Cycle

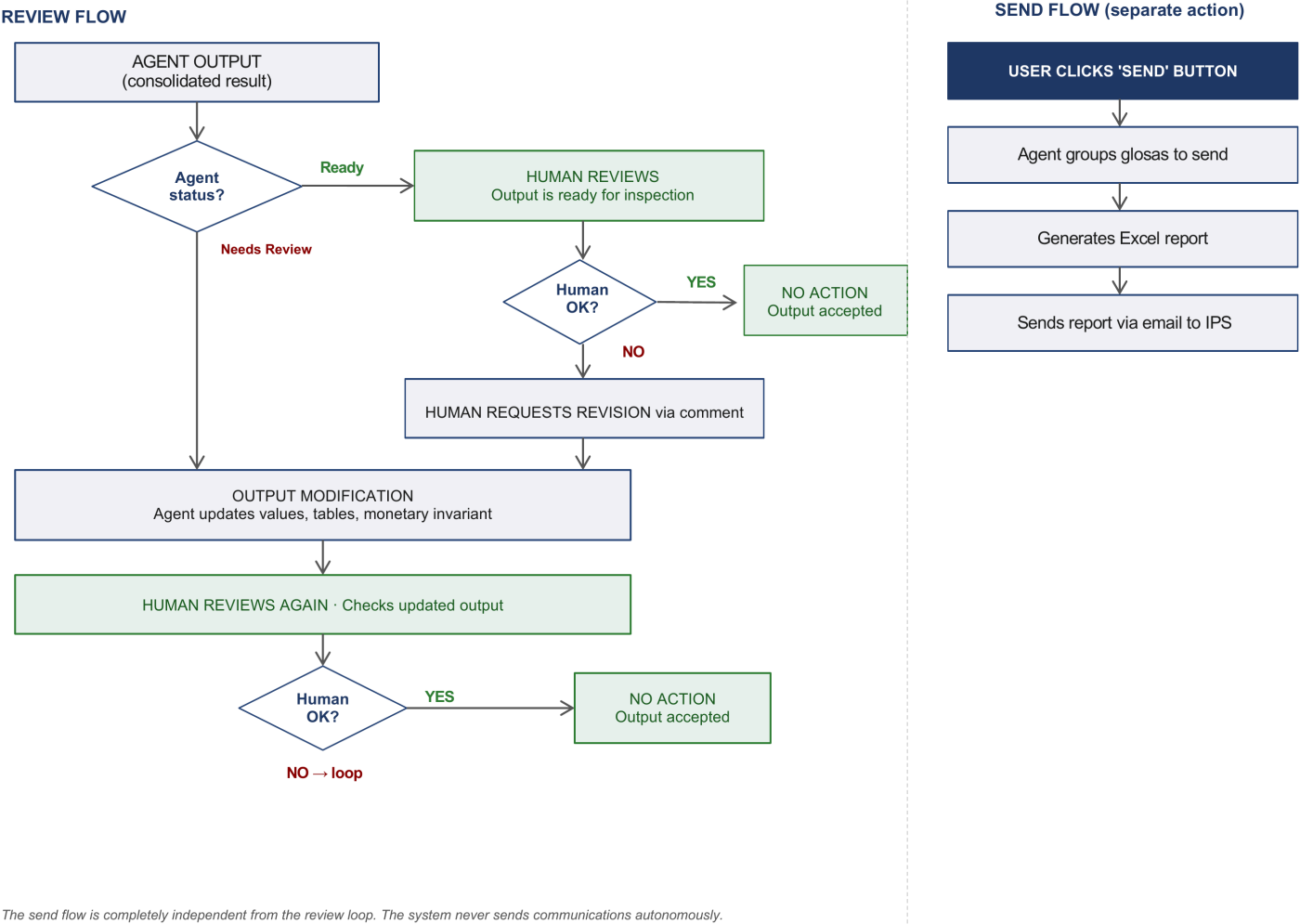


Figure 2

Auditor-supervised correction cycle (review flow) and delivery flow (independent action). In the review flow, the consolidated result enters human review; the auditor may accept it or request corrections via comments iteratively until the result is approved. The delivery flow is completely independent: the user initiates it through an explicit action, the system groups the objections, generates a report, and sends it by email to the provider.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Salmonasupplementary1.docx](#)