

NEUROSYMBOLIC AI SYSTEM FOR AUTOMATED MEDICAL CODING AND RISK ADJUSTMENT

FIELD OF INVENTION

[0001] The present disclosure relates to artificial intelligence systems for healthcare data processing, and more particularly to a neurosymbolic AI system that performs automated medical coding and risk adjustment through member-wise chart review using autonomous AI modules that combine neural networks with structured reasoning for hierarchical condition category (HCC) code identification and validation.

BACKGROUND

[0002] Healthcare organizations increasingly rely on accurate medical coding and risk adjustment processes to ensure proper reimbursement and compliance with regulatory requirements. Risk adjustment systems, particularly those utilizing Hierarchical Condition Categories (HCC), serve as mechanisms to predict future healthcare costs by translating patient medical histories into standardized codes that reflect disease burden and expected resource utilization.

[0003] Traditional medical coding workflows involve manual review of patient electronic medical records by trained human coders who examine clinical documentation to identify appropriate diagnosis codes. These manual processes are time-consuming, labor-intensive, and subject to human error and inconsistency. The conventional approach typically follows a Date of Service (DOS)-wise methodology, where coders review individual patient encounters in isolation, often missing the broader longitudinal context of a patient's health status across multiple visits and time periods.

[0004] Conventional solutions have attempted to address these challenges through various automated coding systems and natural language processing tools. However, these approaches often suffer from limitations including inadequate contextual understanding, fragmented analysis that fails to capture comprehensive patient health profiles, and insufficient integration of structured reasoning with pattern recognition capabilities. Many existing systems rely on simple keyword matching or basic machine learning models that lack the sophistication to handle complex medical terminology and clinical reasoning patterns.

[0005] US Patent Application Publication No. US-20190189253-A1 discloses a medical condition verification system that receives patient electronic medical record data and performs cognitive analysis to identify evidential data supportive of medical conditions. The system generates risk measures based on machine learned relationships of medical factors and provides outputs representing the risk of patients having associated medical conditions. However, this approach focuses primarily on verification of existing medical codes rather than comprehensive automated coding, and lacks the integrated neurosymbolic architecture that combines neural networks with structured reasoning for enhanced accuracy and explainability.

[0006] US Patent Application Publication No. US-20200381090-A1 describes methods for generating patient context vectors as low-dimensional representations of patient medical context using deep learning networks and multi-task learning. The system combines ICD codes and clinical text to create patient context vectors for disease prediction and risk assessment. Nevertheless, this approach is limited to vector-based representations and does not provide the comprehensive member-wise audit capabilities or the structured reasoning components that enable systematic validation and evidence extraction for regulatory compliance.

[0007] Keeping in view the challenges associated with the state of art, there is a need for a neurosymbolic artificial intelligence system that performs automated medical coding through member-wise chart review, combining neural network pattern recognition with structured reasoning capabilities to analyze complete patient records longitudinally. Such a solution would

provide enhanced accuracy in HCC code identification, automated evidence validation, and comprehensive risk adjustment processing while reducing manual workload and improving coding consistency across healthcare organizations.

OBJECTIVES

[0008] The primary objective of the present invention is to provide a neurosymbolic AI system that automates medical coding and risk adjustment through comprehensive member-wise chart review, combining neural network pattern recognition with structured reasoning to improve accuracy and reduce manual workload in healthcare coding processes.

[0009] Another objective of the present invention is to shift from conventional date-of-service-wise review methodologies to a longitudinal analysis approach that may capture complete patient health status across multiple encounters, enabling detection of chronic conditions that may be missed in fragmented encounter-based coding methods.

[0010] Another objective of the present invention is to provide a multi-agent pipeline architecture comprising specialized autonomous modules that may perform distinct functions including encounter segmentation, compliance validation, clinical data extraction, neurosymbolic coding, MEAT evidence validation, hallucination prevention, and member-wise synthesis for comprehensive medical coding automation.

[0011] Another objective of the present invention is to implement automated evidence validation according to Monitoring, Evaluation, Assessment, and Treatment (MEAT) criteria that ensure all recommended codes meet regulatory requirements with supporting clinical documentation for compliance purposes.

[0012] Another objective of the present invention is to provide confidence scoring algorithms that may incorporate multiple weighted parameters including frequency of documentation,

evidence context, MEAT evidence strength, code specificity, and condition type to enable automated triage decisions and hierarchical suppression logic.

[0013] Another objective of the present invention is to reduce cognitive fatigue among human coders by filtering routine coding decisions through automated processing that may allow coding professionals to focus attention on complex cases requiring specialized judgment rather than processing high volumes of straightforward coding decisions.

[0014] Yet another objective of the present invention is to facilitate detection of under-coded chronic conditions documented across multiple visits through comprehensive longitudinal analysis that may improve risk adjustment accuracy and ensure proper reimbursement for healthcare organizations.

SUMMARY

[0015] The present invention relates to a neurosymbolic AI system for automated medical coding and risk adjustment that comprises a multi-agent pipeline architecture configured to process patient medical records, an encounter segmentation agent that parses longitudinal patient charts and identifies discrete clinical encounters by date of service, a compliance check agent that validates encounters against CMS guidelines for HCC coding eligibility, a clinical data extraction agent that extracts medically relevant information based on configurable rules, a neurosymbolic coding agent that combines neural network pattern recognition with structured reasoning from large knowledge models to assign ICD-10 codes, a MEAT validation agent that identifies supporting evidence according to Monitoring, Evaluation, Assessment, and Treatment criteria, a hallucination validator agent that cross-references recommendations against a medical knowledge graph, and a member-wise synthesis module that aggregates findings across all patient encounters to generate HCC code recommendations with confidence scores.

[0016] The system operates by ingesting longitudinal patient charts and claims data, segmenting annual patient records into individual clinical encounters, processing each encounter through the

multi-agent analysis engine to extract clinical information and assign appropriate ICD-10 codes with supporting evidence, performing member-wise synthesis to consolidate findings across all patient visits within a given period, calculating confidence scores based on multiple weighted parameters including frequency of documentation, evidence context, MEAT evidence strength, code specificity, and condition type, and applying decision logic to categorize codes into auto-accept buckets or human validation queues based on confidence thresholds and hierarchical suppression rules.

[0017] The invention provides enhanced accuracy in medical coding through comprehensive longitudinal analysis rather than fragmented encounter-based reviews, reduces manual workload by automating a substantial portion of coding decisions while maintaining high precision, improves coding consistency by applying systematic rules and reasoning across complete patient records, enables detection of under-coded chronic conditions that may be missed in traditional date-of-service reviews, facilitates compliance through automated evidence validation and structured documentation, and reduces cognitive fatigue among human coders by filtering routine decisions and allowing focus on complex cases requiring expert judgment.

BRIEF DESCRIPTION OF FIGURES

[0018] Non-limiting and non-exhaustive examples are described with reference to the following figures.

[0019] Figure 1(a-b) illustrates a workflow diagram of a neurosymbolic auditing framework pipeline for medical coding.

[0020] FIG. 2 illustrates a block diagram of an automated risk adjustment coding system.

DETAILED DESCRIPTION

[0021] The following detailed description and embodiments set forth herein below are merely exemplary out of the wide variety and arrangement of instructions which can be employed with the present invention. The present invention may be embodied in other specific forms without

departing from the spirit or essential characteristics thereof. All the features disclosed in this specification may be replaced by similar other or alternative features performing similar or same or equivalent purposes. Thus, unless expressly stated otherwise, they all are within the scope of the present invention.

[0022] Accordingly, those of ordinary skill in the art will recognize that various changes and modifications of the embodiments described herein can be made without departing from the scope of the invention. In addition, descriptions of well-known functions and constructions are omitted for clarity and conciseness.

[0023] The terms and words used in the following description and claims are not limited to the bibliographical meanings but are merely used to enable a clear and consistent understanding of the invention. Accordingly, it should be apparent to those skilled in the art that the following description of exemplary embodiments of the present invention are provided for illustration purpose only and not for the purpose of limiting the invention.

[0024] It is to be understood that the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise.

[0025] It should be emphasized that the term “comprises/comprising” when used in this specification is taken to specify the presence of stated features, integers, steps, or components but does not preclude the presence or addition of one or more other features, integers, steps, components, or groups thereof.

[0026] The present invention addresses challenges in risk adjustment coding by providing a neurosymbolic auditing framework for automated Hierarchical Condition Category (HCC) coding in healthcare systems. Risk adjustment systems may be used to ensure that insurance providers receive appropriate compensation based on the expected costs of their members, with health plans covering sicker, higher-cost patients receiving higher payments than plans covering healthier populations. The system may rely on accurately translating a patient's medical history into ICD-10 codes, which then map to Hierarchical Condition Categories (HCC).

[0027] In an embodiment, the present invention provides a neurosymbolic auditing framework that combines the pattern-recognition capabilities of neural networks with the structured reasoning of large knowledge models to perform comprehensive member-wise chart review for risk adjustment HCC coding. The neurosymbolic auditing framework may leverage an agentic AI approach that operates autonomously to achieve specified goals by executing a sequence of steps, using tools, and adapting strategies based on processed information. Agentic AI may refer to systems that can operate autonomously to achieve a specified goal by executing a sequence of steps, using tools, and adapting strategies based on the information processed. The framework may shift analysis from traditional Date of Service (DOS)-wise review to a member-wise audit approach, where the system consolidates and analyzes an entire patient record for a given period to identify chronic conditions and recommend HCC with greater accuracy and context.

[0028] FIG. 1(a) along with FIG 1(b) illustrates a flowchart for a neurosymbolic auditing framework method 100 for automated risk adjustment coding. The method 100 begins with step 102, where the system receives longitudinal patient chart data and corresponding claims data. The process then proceeds to step 104, which involves segmenting the longitudinal patient chart data into discrete clinical encounters based on date of service.

[0029] The method 100 continues to step 106, where each discrete clinical encounter is processed through a multi-agent analysis engine comprising a plurality of specialized AI agents that sequentially analyze the clinical encounter to extract clinical data and assign ICD-10 codes. Following this analysis, the process moves to step 108, which performs member-wise synthesis by aggregating findings from all encounters to generate a holistic patient view.

[0030] The method 100 then advances to step 110, where assigned ICD-10 codes are filtered to retain only codes that map to Hierarchical Condition Categories (HCC). Subsequently, the process proceeds to step 112, which calculates a confidence score for each retained HCC based on multiple weighted parameters including frequency of documentation across encounters, condition evidence context, and MEAT evidence strength.

[0031] The method 100 concludes with step 114, where coding recommendations are generated by categorizing each HCC into either an auto-accept bucket for codes with high confidence scores or a human validation queue for codes with low or medium confidence scores. The flowchart demonstrates a sequential process flow where each step builds upon the previous

step's output, transforming raw longitudinal patient data into structured coding recommendations through automated analysis and intelligent categorization.

[0032] The present invention also relates to a method for implementing automated risk adjustment coding in a data processing system. The method comprises the following steps: The neurosymbolic auditing framework may receive longitudinal patient chart data and corresponding claims data for a patient over a given period. The system may process annual patient charts that are typically delivered as single large PDF documents alongside corresponding claims data. The framework may handle data from multiple distinct anonymized datasets representing a wide variety of chart formats, clinical complexities, and data quality levels to ensure robust and generalizable performance across different clinical environments.

[0033] The neurosymbolic auditing framework may segment the longitudinal patient chart data into discrete clinical encounters based on DOS. This segmentation process may involve parsing the entire patient chart to identify the DOS for each distinct encounter and establish precise document boundaries for each visit. The segmentation may effectively deconstruct a single, year-long PDF into multiple, visit-specific documents while extracting and validating demographic details from each visit to detect anomalies and ensure the integrity of the patient record.

[0034] The neurosymbolic auditing framework may process each discrete clinical encounter through a multi-agent analysis engine comprising a plurality of specialized AI agents that sequentially analyze the clinical encounter to extract clinical data and assign ICD-10 codes. Each visit-specific document may be processed through the core analysis engine where specialized agents work in sequence to transform raw, unstructured patient data into structured, auditable coding recommendations.

[0035] The neurosymbolic auditing framework may perform member-wise synthesis by aggregating findings from all discrete clinical encounters for the patient to generate a holistic patient view. This synthesis process may consolidate all findings from individual visits into a comprehensive member-wise perspective, allowing the system to understand the complete patient journey and identify chronic conditions that may be documented across multiple encounters over time.

[0036] The neurosymbolic auditing framework may filter the assigned ICD-10 codes to retain only codes that map to Hierarchical Condition Categories (HCC). The filtering process may focus on codes that map to V28 HCC model or RxHCC categories, with the V28 HCC model covering

8,183 ICD-to-HCC mappings with 7,903 unique ICDs and the RxHCC model covering 5,299 mappings with 5,159 unique ICDs to ensure comprehensive coverage of relevant diagnostic categories.

[0037] The neurosymbolic auditing framework may calculate a confidence score for each retained HCC based on multiple weighted parameters including frequency of documentation across encounters, condition evidence context, and MEAT evidence strength. The confidence scoring may integrate additional parameters such as code specificity based on similarity between documented evidence and official ICD-10 code descriptions, and condition type weighting that assigns higher weights to chronic conditions compared to acute conditions.

[0038] FIG. 2 illustrates a block diagram of an automated risk adjustment coding system 202. The automated risk adjustment coding system 202 includes a data processing system 204, a neurosymbolic auditing framework 210, a medical knowledge graph 230, a demographic validation component 232, and hierarchical suppression logic 234. The data processing system 204 comprises a processor 206 and memory 208. The processor 206 executes instructions stored in the memory 208 to control the operation of the system components. The data processing system 204 is connected to the neurosymbolic auditing framework 210, which serves as the primary processing component for automated HCC coding operations. The neurosymbolic auditing framework 210 contains a multi-agent analysis engine 212, a member-wise synthesis module 224, a confidence scoring system 226, and an output categorization system 228. The multi-agent analysis engine 212 includes five specialized AI agents that operate sequentially: a compliance check agent 214, a clinical data extraction agent 216, a neurosymbolic coding agent 218, a MEAT validation agent 220, and a hallucination validator agent 222.

[0039] The multi-agent analysis engine 212 connects to the member-wise synthesis module 224, which aggregates findings from all discrete clinical encounters to generate a holistic patient view. The member-wise synthesis module 224 connects to the confidence scoring system 226, which calculates confidence scores for each retained HCC code based on multiple weighted parameters. The confidence scoring system 226 connects to the output categorization system 228, which implements a triage system to categorize codes into auto-accept buckets or human validation queues. The hallucination validator agent 222 connects to the medical knowledge graph 230, which provides cross-referencing capabilities for validating AI-generated recommendations against established medical knowledge. The neurosymbolic auditing framework 210 connects to the

demographic validation component 232, which cross-verifies patient information across encounters to detect anomalies. The output categorization system 228 connects to the hierarchical suppression logic 234, which automatically suppresses redundant codes and filters lower-hierarchy HCC codes to maximize processing efficiency.

[0040] The system for automated risk adjustment coding may comprise several interconnected components that work together to implement the neurosymbolic auditing framework. The system architecture may be designed to process longitudinal patient data through a series of specialized components that transform raw clinical information into structured coding recommendations.

[0041] (a) A data processing system may include at least one processor and at least one memory comprising instructions that, when executed by the at least one processor, cause the at least one processor to implement the neurosymbolic auditing framework. The data processing system may serve as the foundational computing infrastructure that coordinates all system operations and may be communicatively connected to data sources providing longitudinal patient chart data and corresponding claims data. The processor may execute the instructions stored in the memory to control the operation of all other system components and may manage data flow between the various specialized agents and processing modules.

[0042] (b) A neurosymbolic auditing framework may be implemented by the data processing system and may be configured to receive longitudinal patient chart data and corresponding claims data for a patient over a given period. The neurosymbolic auditing framework may be communicatively connected to the data processing system and may coordinate the overall workflow from data ingestion through final code recommendations. The framework may combine pattern-recognition capabilities of neural networks with structured reasoning of large knowledge models to perform comprehensive analysis of patient records.

[0043] (c) A multi-agent analysis engine may be integrated within the neurosymbolic auditing framework and may comprise a plurality of specialized AI agents that sequentially analyze clinical encounters. The multi-agent analysis engine may be communicatively connected to the neurosymbolic auditing framework and may receive segmented clinical encounter data for processing. The engine may coordinate the sequential operation of specialized agents including a compliance check agent, a clinical data extraction agent, a neurosymbolic coding agent, a MEAT validation agent, and a hallucination validator agent, with each agent passing processed data to the subsequent agent in the pipeline.

[0044] (d) A demographic validation component may be integrated within the segmentation process and may cross-verify patient information across all encounters to detect anomalies such as multiple patients within a single chart. The demographic validation component may be communicatively connected to the encounter segmentation functionality and may extract demographic details from each visit to ensure data integrity. The component may flag inconsistencies in patient identification information and may prevent processing of corrupted or mixed patient records.

[0045] (e) A medical knowledge graph may be integrated within the system and may provide cross-referencing capabilities for recommended codes and their supporting evidence. The medical knowledge graph may be communicatively connected to the hallucination validator agent and may contain structured medical knowledge used to verify the plausibility of coding recommendations. The knowledge graph may store relationships between medical conditions, treatments, and diagnostic codes to enable validation of AI-generated recommendations against established medical knowledge.

[0046] (f) An output categorization system may be communicatively connected to the neurosymbolic auditing framework and may implement a triage system that categorizes final output into auto-accept buckets and human validation queues. The output categorization system may receive confidence scores and data source classifications from the framework and may apply decision logic to route codes appropriately. The system may automatically categorize codes with high confidence scores into auto-accept buckets while directing codes with low or medium confidence scores, or those originating from orphan charts or orphan claims, to human validation queues regardless of confidence score.

[0047] The neurosymbolic auditing framework may be configured to segment the longitudinal patient chart data into discrete clinical encounters based on DOS, with the segmentation process involving parsing of complete patient charts to identify distinct encounters and establish document boundaries. The framework may process each discrete clinical encounter through the multi-agent analysis engine to extract clinical data and assign ICD-10 codes through sequential agent operations. The framework may perform member-wise synthesis by aggregating findings from all discrete clinical encounters to generate a holistic patient view that captures chronic conditions documented across multiple visits. The framework may filter assigned ICD-10 codes to retain only codes that map to Hierarchical Condition Categories and may calculate confidence scores for each

retained HCC based on multiple weighted parameters. The framework may generate coding recommendations by categorizing each HCC based on the calculated confidence score through the output categorization system. The system may further comprise integration interfaces configured to communicate with existing healthcare information systems to receive patient data and transmit coding recommendations.

[0048] The advantages of the present invention may include substantial improvements in coding accuracy and efficiency compared to traditional manual processes. The member-wise approach may enable identification of chronic conditions that span multiple encounters, reducing missed diagnoses that occur in fragmented DOS-wise reviews. The automated confidence scoring and triage system may reduce manual workload by enabling straight-through processing of high-confidence codes while focusing human expertise on complex cases requiring nuanced judgment. The system may achieve high precision in auto-accepted codes while maintaining comprehensive coverage with minimal manual additions, thereby reducing coder fatigue and improving overall accuracy in risk adjustment operations.

[0049] The advantages of the present invention may include substantial improvements in coding accuracy and efficiency compared to traditional manual processes. The member-wise approach may enable identification of chronic conditions that span multiple encounters, reducing missed diagnoses that occur in fragmented DOS-wise reviews. The automated confidence scoring and triage framework may reduce manual workload by enabling straight-through processing of high-confidence codes while focusing human expertise on complex cases requiring nuanced judgment. The framework may achieve high precision in auto-accepted codes while maintaining comprehensive coverage with minimal manual additions, thereby reducing coder fatigue and improving overall accuracy in risk adjustment operations.

[0050] Further, the experiment is conducted on four distinct, anonymized datasets sourced from real-world clinical operations. These data sets represent a wide variety of chart formats, clinical complexities, and data quality levels, thereby ensuring a robust and generalizable evaluation of the system's performance. For each member in the datasets, the available information included one full year of clinical charts and the corresponding claims data. A set of correct and auditable HCC codes was established for each member. A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made without

departing from the spirit and scope of the disclosure. Accordingly, other implementations are within the scope of the following claims.

EXPERIMENTAL METHODOLOGY

[0051] The evaluation for each member in the datasets may follow a systematic data processing workflow that applies the agents in a sequential manner. The evaluation process may comprise four distinct phases: (1) Ingestion and Segmentation, where for each member, the complete one-year clinical chart and claims data may be ingested, and the encounter segmentation agent may parse the longitudinal chart to identify and isolate each discrete clinical encounter into a separate, visit-specific document; (2) Encounter-Level Analysis, where each segmented encounter may be subsequently processed by the multi-agent core analysis engine through sequential application of the compliance check agent, the clinical data extraction agent, the neurosymbolic coding agent, the MEAT validation agent, and the hallucination validator agent to produce a set of validated findings for each visit; (3) Member-Wise Synthesis and Scoring, where following the analysis of all encounters for a member, the framework may perform member-wise synthesis by aggregating all findings and retaining only ICD-10 codes mapping to V28 HCC or RxHCC categories, with the system calculating a final confidence score for each unique code based on the consolidated evidence; and (4) Triage for Evaluation, where the system may apply the decision framework to triage the scored recommendations by categorizing each code into either the auto-accept bucket or the human validation queue, with this final categorized output forming the basis for performance evaluation.

[0052] The final output from the experimental workflow may be categorized for analysis into several distinct groups, allowing for a granular assessment of the system's performance. The framework may categorize as-

[0053] Total System Recommendations as the complete set of unique ICD-10 codes suggested by the framework for a member, prior to any human interaction.

[0054] Auto-Accepted Recommendations may comprise a subset of recommendations that the system identified with HIGH confidence.

[0055] Recommendations Flagged for Human Review may include codes that the system identified with LOW or MEDIUM confidence, or those originating from orphan charts/claims, with this queue including Unclaimed Codes representing potential under-coding and Overclaimed Codes representing potential over-coding.

[0056] Manual Code Additions may comprise any valid codes that were missed entirely by the system and subsequently added by the human coder during their final review.

RESULTS

[0057] To provide a rigorous quantitative evaluation, the framework may define five key metrics. Let M be the set of all members in the evaluation dataset. For a given member $m \in M$, the following sets of codes may be defined:

- $S_{AA}^{(m)}$ as the set of codes placed in the Auto-Accept bucket by the system;
- $S_U^{(m)}$ as the set of Unclaimed (Net New) codes recommended by the system for human review;
- $S_O^{(m)}$ as the set of Overclaimed codes flagged by the system for human review;
- $S_{Sys}^{(m)}$ as the set of all unique codes recommended by the system ($S_{AA}^{(m)} \cup S_U^{(m)}$);
- $S_{Man}^{(m)}$ as the set of valid codes added manually by a human coder that were missed by the system; and $S_{Final}^{(m)}$ as the final set of all validated, correct codes for the member (the "gold standard").

[0058] A validation function, $V(C,m)$, may also be defined, which returns 1 if code C is in the gold standard set $S_{Final}^{(m)}$, and 0 otherwise. The overall performance metrics may then be calculated by summing the counts across all members:

1. Auto-Accept Precision (P_{AA}): May measure the accuracy of the codes placed in the auto-accept bucket.

$$P_{AA} = \frac{\sum_{m \in M} \sum_{C \in S_{AA}^{(m)}} V(C,m)}{\sum_{m \in M} |S_{AA}^{(m)}|} \times 100\%$$

2. Net New Code Discovery Rate (D_{NN}): May measure the system's effectiveness at identifying valid, under-claimed codes.

$$D_{NN} = \frac{\sum_{m \in M} \sum_{C \in S_U^{(m)}} V(C,m)}{\sum_{m \in M} |S_U^{(m)}|} \times 100\%$$

3. Overclaimed Code Rejection Accuracy (A_{OR}): May measure the system's ability to correctly flag codes from claims that lack sufficient clinical evidence.

$$A_{OR} = \frac{\sum_{m \in M} \sum_{C \in S_m^{(m)}} (1 - V(C, m))}{\sum_{m \in M} |S_m^{(m)}|} \times 100\%$$

4. Autonomous Coding Rate (R_{Auto}): May represent the proportion of the total recommended codes that were handled automatically.

$$R_{Auto} = \frac{\sum_{m \in M} |S_{AA}^{(m)}|}{\sum_{m \in M} |S_{Dys}^{(m)}|} \times 100\%$$

5. Manual Addition Rate (R_{Man}): May measure the dependency on human intervention for codes the system missed entirely.

$$R_{Man} = \frac{\sum_{m \in M} |S_{Man}^{(m)}|}{\sum_{m \in M} |S_{Final}^{(m)}|} \times 100\%$$

[0059] The results presented in Table [XX] may strongly support the initial hypothesis. The high Auto-Accept Precision of 98.29% may demonstrate the robustness of the multi-parameter confidence scoring, confirming that the system may reliably triage a significant portion of codes for straight-through processing. This may directly contribute to the Autonomous Coding Rate of 82.19%, which may signify a substantial reduction in the manual workload for coding teams.

Furthermore, the framework may show strong performance in addressing both under-coding and over-coding. The Net New Code Discovery Rate of [XX.X]% may indicate that the member-wise analysis is highly effective at identifying valid chronic conditions that may be often missed in fragmented, DOS-wise reviews. Conversely, the high Overclaimed Code Rejection Accuracy of [XX.X]% may highlight the system's ability to enforce compliance by flagging claimed codes that lack the necessary MEAT evidence in the clinical documentation.

[0060] Finally, the low Manual Addition Rate of just 1.72% may suggest that the system exhibits high recall, with very few valid codes being missed entirely by the automated analysis. This comprehensive performance may underscore the value of the neurosymbolic, member-wise approach in creating a more accurate, efficient, and compliant risk adjustment process.

[0061] Table 1- A detailed breakdown of the system's output, categorizing all code recommendations across the four datasets and in aggregate. The table presents the raw counts for codes placed in the 'Auto-Accept' bucket, those flagged for 'Required Validation' (including Unclaimed and Overclaimed), and codes missed by the system and subsequently 'Added by Coder'. The results highlight the high volume of automated decisions (877 codes in the Auto-Accept bucket overall) and the low number of manual additions (only 17 across all datasets), indicating the framework's comprehensive coverage and scale.

Category	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Overall
Total Suggested Codes	330	196	150	381	1067
Auto-Accept Bucket	273	146	120	338	877
Accepted	268	142	117	335	862
Rejected	5	4	3	3	15
Required Validation Bucket	85	70	30	53	238
Unclaimed Accepted	32	29	18	31	110
Unclaimed Rejected	25	21	12	22	80
Overclaimed Accepted	1	0	1	0	2
Overclaimed Rejected	27	18	21	18	84
Added by Coder (System Miss)	3	2	3	9	17

[0062] The quantitative results from the experiment, evaluating the neurosymbolic auditing framework against the gold standard across four distinct datasets. The findings may be organized to demonstrate the framework's performance in terms of recommendation accuracy, its significant impact on coder efficiency through the member-wise paradigm, and a comparative analysis against established human coder reliability benchmarks.

QUANTITATIVE PERFORMANCE ANALYSIS

[0063] The framework's performance may be measured by comparing its output against the gold standard. A detailed breakdown of the raw code counts, including aggregated totals, may be presented in Table 1.

[0064] From these counts, the key performance metrics may be calculated, which are summarized in Table 2.

[0065] The overall results may highlight the framework's robustness. An Auto-Accept Precision of 98.29% across 877 high-confidence decisions may validate the reliability of the confidence scoring model. This may enable a high Autonomous Coding Rate of 82.19%, signifying that the majority of the analytical workload can be automated. Furthermore, the extremely low overall Manual Addition Rate of 1.72% may demonstrate the system's high recall and comprehensive coverage.

Table 2. Summary of the key performance metrics, calculated from the raw data presented in Table 1. This table quantifies the framework's effectiveness across all four datasets and in aggregate. The results highlight the system's high reliability, evidenced by an overall Auto-Accept Precision of 98.29%, and its significant impact on workflow, shown by an 82.19% Autonomous Coding Rate. The low Manual Addition Rate of 1.72% further underscores the system's high recall, confirming that it rarely misses valid codes.

Performance Metric	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Overall
Auto-Accept Precision (P_{AA})	98.17%	97.26%	97.50%	99.11%	98.29%
Net New Code Discovery Rate (D_{NN})	56.14%	58.00%	60.00%	58.49%	57.89%
Overclaimed Rejection Accuracy (A_{OR})	98.43%	100.00%	95.45%	100.00%	97.67%
Autonomous Coding Rate (R_{Aut})	82.73%	74.49%	80.00%	86.45%	82.19%
Manual Addition Rate (R_{Man})	0.99%	1.16%	2.16%	2.40%	1.72%

EFFICIENCY GAINS: MEMBER-WISE VS. DOS-WISE CURATION

[0066] A central hypothesis of this work may be that a member-wise audit is not only more accurate but also significantly more efficient. To test this, a comparative case study may be conducted on a representative member, tasking a coder with validating the output of a traditional DOS-wise presentation versus the intelligent member-wise presentation. The results may be summarized in Table 3.

Table 3. A case study comparison illustrating the efficiency gains of the member-wise curation paradigm over a traditional DOS-wise review. For an identical set of 16 system-suggested codes for a single member, this table contrasts the number of items presented to a human coder and the total time required for validation. The member-wise approach, by intelligently suppressing 50%

of the codes (8 of 16) that were redundant or hierarchically inferior, reduced the coder's manual effort and total curation time by 81%. This demonstrates the significant impact of the framework's synthesis and filtering logic on human productivity.

Item	DOS-wise Coding	Member-wise Coding	Comment
Total codes suggested by system	16	16	Total findings are identical.
Codes presented to coder for curation	16	6	62% reduction in codes requiring review.
Codes manually curated by coder	16	3	81% reduction in manual interaction.
Codes automatically suppressed by system	0	8	50% of codes intelligently filtered by the system.
Time taken by coder (minutes)	21	4	81% reduction in total curation time.

[0067] The efficiency gains may be substantial. By aggregating evidence on a member-wise basis, the system's suppression logic may intelligently filter out redundant, duplicative, or lower-hierarchy HCC. This may reduce the number of codes presented to the coder by 62% and the total time for curation by 81%, confirming that the member-wise paradigm, powered by neurosymbolic AI, may dramatically reduce the cognitive load and manual effort required for validation.

COMPARATIVE ANALYSIS AGAINST HUMAN CODER RELIABILITY

To contextualize the system's performance, the Auto-Accept Precision may be compared to benchmarks of human inter-coder reliability. A study on the consistency of medical coding (PMCID: PMC7299481) may have found that the agreement between human coders for the same condition was approximately 82-86%. The framework's overall Auto-Accept Precision of 98.29% may significantly exceed this benchmark, suggesting that the system operates not only at the level of a human expert but with a higher degree of consistency than may typically be found between two independent human coders.

[0068] Further, the system explicitly handles the common scenarios of misalignment between clinical charts and billing claims.

1. Matched Pair: The ideal scenario where a patient visit chart has a corresponding claim date. These are processed for potential auto-acceptance.
2. Orphan Chart: An encounter for which a clinical chart exists but no corresponding claim has been submitted.
3. Orphan Claim: An encounter for which a claim has been submitted but for which no supporting clinical documentation has been provided.

Further, codes originating from orphan charts or orphan claims are automatically routed for human validation. For matched pairs, the decision is based on the calculated confidence score. Codes with a HIGH score are designated for the Auto-Accept bucket, requiring no further human review. Codes with LOW or MEDIUM scores are flagged for Human Validation. Hierarchical Suppression Logic: To maximize efficiency and prevent redundant coding, a suppression rule is applied. When any code is designated for auto-acceptance, all other codes that map to the same HCC or RxHCC or to a child HCC/RxHCC in the hierarchy, are automatically suppressed for that member. This ensures that only the most confident and highest-value codes are put forward.

WE CLAIM:

1. A method (100), in a data processing system to implement a neurosymbolic auditing framework, the method comprising:
 - receiving, by the neurosymbolic auditing framework (210), longitudinal patient chart data and corresponding claims data for a patient over a given period;
 - segmenting, by the neurosymbolic auditing framework (210), the longitudinal patient chart data into discrete clinical encounters based on DOS;
 - processing, by the neurosymbolic auditing framework (210), each discrete clinical encounter through a multi-agent analysis engine (212) comprising a plurality of specialized AI agents that sequentially analyze the clinical encounter to extract clinical data and assign ICD-10 codes;
 - performing, by the neurosymbolic auditing framework (210), member-wise synthesis by aggregating findings from all discrete clinical encounters for the patient to generate a holistic patient view;
 - filtering, by the neurosymbolic auditing framework (210), the assigned ICD-10 codes to retain only codes that map to Hierarchical Condition Categories (HCC);
 - calculating, by the neurosymbolic auditing framework (210), a confidence score for each retained HCC based on multiple weighted parameters including frequency of documentation across encounters, condition evidence context, and MEAT evidence strength; and
 - generating, by the neurosymbolic auditing framework (210), coding recommendations by categorizing each HCC into either an auto-accept bucket for codes with high confidence scores or a human validation queue for codes with low or medium confidence scores.
2. The method of claim 1, wherein the plurality of specialized AI agents comprises:
 - a compliance check agent (214) that verifies each discrete clinical encounter meets regulatory guidelines for HCC coding eligibility;
 - a clinical data extraction agent (216) that extracts medically relevant information based on configurable rules derived from regulatory guidelines;
 - a neurosymbolic coding agent (218) that analyzes extracted clinical data and assigns ICD-10 codes using a large knowledge model;

- a MEAT validation agent (220) that extracts supporting evidence according to MEAT criteria; and
 - a hallucination validator agent (222) that cross-references recommended codes against a medical knowledge graph (230).
3. The method of claim 2, wherein the compliance check agent (214) filters out non-compliant interactions including telephonic-only calls and retains only face-to-face visits for further processing.
 4. The method of claim 1, wherein the multiple weighted parameters for calculating the confidence score further comprise:
 - code specificity based on similarity between documented evidence and official ICD-10 code descriptions; and
 - condition type weighting that assigns higher weights to chronic conditions compared to acute conditions.
 5. The method of claim 1, wherein performing member-wise synthesis comprises:
 - mapping the assigned ICD-10 codes to V28 HCC model or RxHCC categories;
 - aggregating all associated data points from across all discrete clinical encounters for each valid ICD-10 code; and
 - applying hierarchical suppression logic that automatically suppresses codes mapping to the same HCC or child HCC when a higher confidence code is designated for auto-acceptance.
 6. The method of claim 5, wherein the hierarchical suppression logic reduces the number of codes requiring human validation by automatically filtering redundant and lower-hierarchy HCC.
 7. The method of claim 1, further comprising handling data discrepancies between the longitudinal patient chart data and the corresponding claims data by categorizing encounters as:
 - matched pairs where longitudinal patient chart data corresponds to claim dates;
 - orphan charts where longitudinal patient chart data exists without corresponding claims; and
 - orphan claims where claims exist without supporting longitudinal patient chart data.
 8. The method of claim 7, wherein codes originating from orphan charts or orphan claims are automatically routed to the human validation queue regardless of confidence score.
 9. A system for automated risk adjustment coding (202), comprising:

a data processing system (204) comprising at least one processor (206) and at least one memory (208) comprising instructions that, when executed by the at least one processor (206), cause the at least one processor (206) to implement:

a neurosymbolic auditing framework (210) configured to:

receive longitudinal patient chart data and corresponding claims data for a patient over a given period,

segment the longitudinal patient chart data into discrete clinical encounters based on date of service,

process each discrete clinical encounter through a multi-agent analysis engine (212) comprising a plurality of specialized AI agents,

perform member-wise synthesis by aggregating findings from all discrete clinical encounters for the patient to generate a holistic patient view,

filter assigned ICD-10 codes to retain only codes that map to Hierarchical Condition Categories (HCC),

calculate a confidence score for each retained HCC based on multiple weighted parameters, and generate coding recommendations by categorizing each HCC based on the confidence score;

a multi-agent analysis engine (212) comprising a compliance check agent (214) that verifies regulatory guidelines compliance, a clinical data extraction agent (216) that extracts medically relevant information, a neurosymbolic coding agent (218) that analyzes data and assigns ICD-10 codes, a MEAT validation agent (220) that extracts supporting evidence per MEAT criteria, and a hallucination validator agent (222) that cross-references codes against medical knowledge;

a demographic validation component (232) that cross-verifies patient information across encounters and detects anomalies;

a medical knowledge graph (230) that provides cross-referencing capabilities and validates AI-generated recommendations;

an output categorization system (228) that implements a triage system to route codes to auto-accept buckets or human validation queues based on confidence scores;

an encounter segmentation process that segments longitudinal data into discrete clinical encounters;

a member-wise synthesis module (224) that aggregates findings across all encounters;
a confidence scoring system (226) that calculates scores based on multiple weighted parameters; and

a hierarchical suppression logic (234) that automatically suppresses redundant codes and filters lower-hierarchy HCC codes.

10. The system of claim 9, wherein the plurality of specialized AI agents comprises:

a compliance check agent (214) that verifies each discrete clinical encounter meets regulatory guidelines for HCC coding eligibility;

a clinical data extraction agent (216) that extracts medically relevant information based on configurable rules derived from regulatory guidelines;

a neurosymbolic coding agent (218) that analyzes extracted clinical data and assigns ICD-10 codes using a large knowledge model;

a MEAT validation agent (220) that extracts supporting evidence according to MEAT criteria; and

a hallucination validator agent (222) that cross-references recommended codes against a medical knowledge graph (230).

11. The system of claim 10, wherein the compliance check agent (214) filters out non-compliant interactions including telephonic-only calls and retains only face-to-face visits for further processing.

12. The system of claim 9, wherein the multiple weighted parameters for calculating the confidence score comprise:

frequency of documentation across encounters;

condition evidence context based on section headers where conditions are documented;

MEAT evidence strength based on type of supporting evidence;

code specificity based on similarity between documented evidence and official ICD-10 code descriptions; and

condition type weighting that assigns higher weights to chronic conditions compared to acute conditions.

13. The system of claim 9, wherein the neurosymbolic auditing framework (210) is further configured to perform member-wise synthesis by:

mapping the assigned ICD-10 codes to V28 HCC model or RxHCC categories;
aggregating all associated data points from across all discrete clinical encounters for each valid ICD-10 code; and

applying hierarchical suppression logic (234) that automatically suppresses codes mapping to the same HCC or child HCC when a higher confidence code is designated for auto-acceptance.

14. The system of claim 13, wherein the hierarchical suppression logic (234) reduces the number of codes requiring human validation by automatically filtering redundant and lower-hierarchy HCC codes.
15. The system of claim 9, wherein the neurosymbolic auditing framework (210) is further configured to handle data discrepancies between the longitudinal patient chart data and the corresponding claims data by categorizing encounters as matched pairs, orphan charts, and orphan claims, and automatically routing codes originating from orphan charts or orphan claims to a human validation queue regardless of confidence score.
16. The system of claim 9, wherein the multi-agent analysis engine (212) is configured to process each discrete clinical encounter through sequential analysis by the plurality of specialized AI agents to extract clinical data and assign ICD-10 codes.
17. The system of claim 9, further comprising a data discrepancy handling module configured to categorize encounters as matched pairs where clinical chart data corresponds to claim dates, orphan charts where clinical chart data exists without corresponding claims, and orphan claims where claims exist without supporting clinical documentation.
18. The system of claim 9, further comprising integration interfaces configured to communicate with existing healthcare information systems to receive patient data and transmit coding recommendations.
19. The system of claim 9, wherein the system is configured to achieve performance metrics including an Auto-Accept Precision rate, an Autonomous Coding Rate, and a Manual Addition Rate for measuring coding accuracy and automation efficiency.

NEUROSYMBOLIC AI SYSTEM FOR AUTOMATED MEDICAL CODING AND RISK ADJUSTMENT

ABSTRACT

The present disclosure provides a method in a data processing system to implement a neurosymbolic auditing framework. The method comprises receiving longitudinal patient chart data and corresponding claims data for a patient, segmenting the data into discrete clinical encounters based on date of service, and processing each encounter through a multi-agent analysis engine comprising specialized AI agents that sequentially analyze the encounter to extract clinical data and assign ICD-10 codes. The method further comprises performing member-wise synthesis by aggregating findings from all encounters, filtering assigned ICD-10 codes to retain codes mapping to Hierarchical Condition Categories (HCCs), calculating confidence scores for each HCC code based on weighted parameters including documentation frequency, condition evidence context, and MEAT evidence strength, and generating coding recommendations by categorizing each HCC code into either an auto-accept bucket for high confidence scores or human validation queue for low or medium confidence scores.

Fig. 1(a)

100

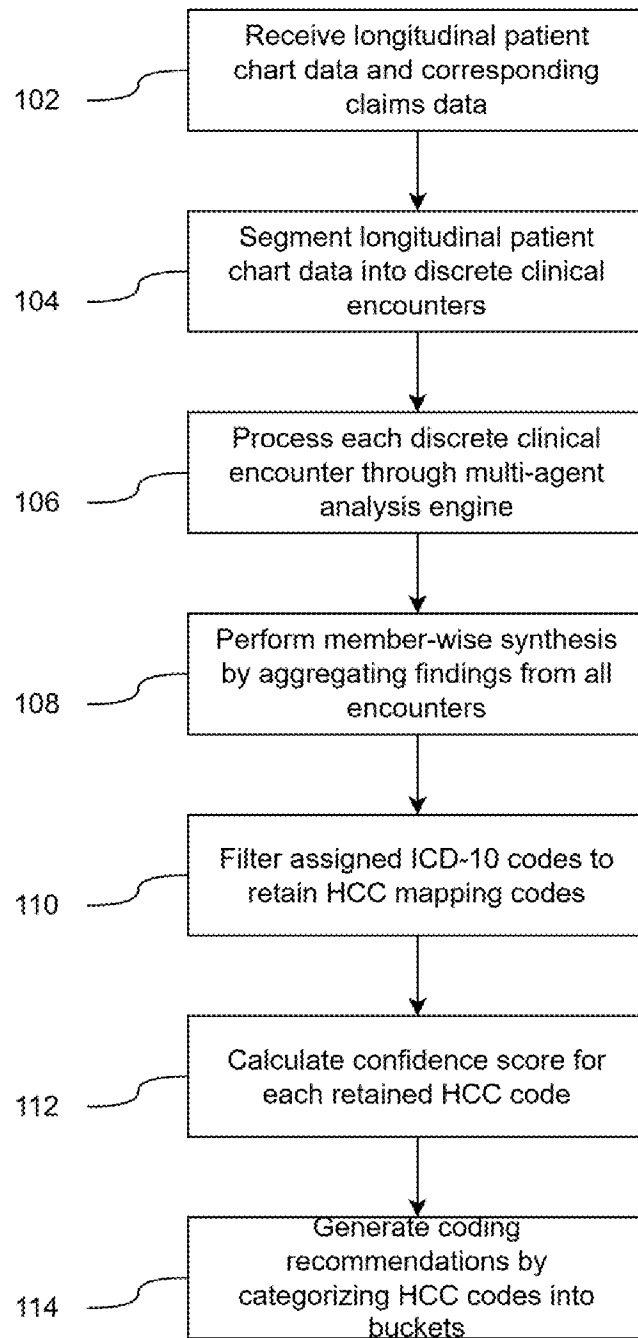


Figure 1(a)

Neurosymbolic Auditing Framework Pipeline

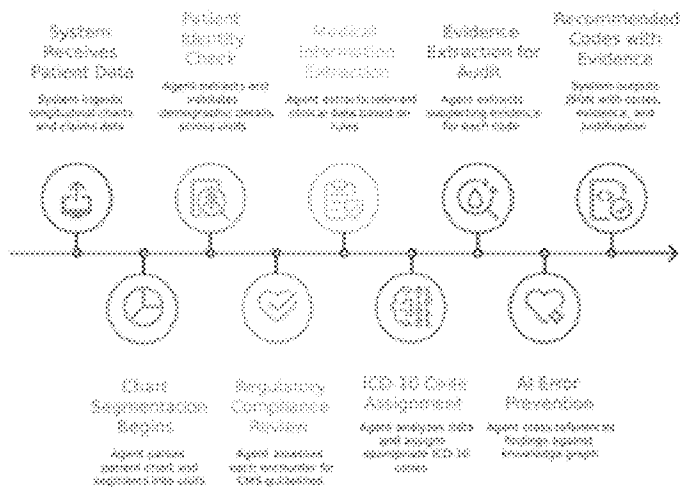


Figure 1(b)

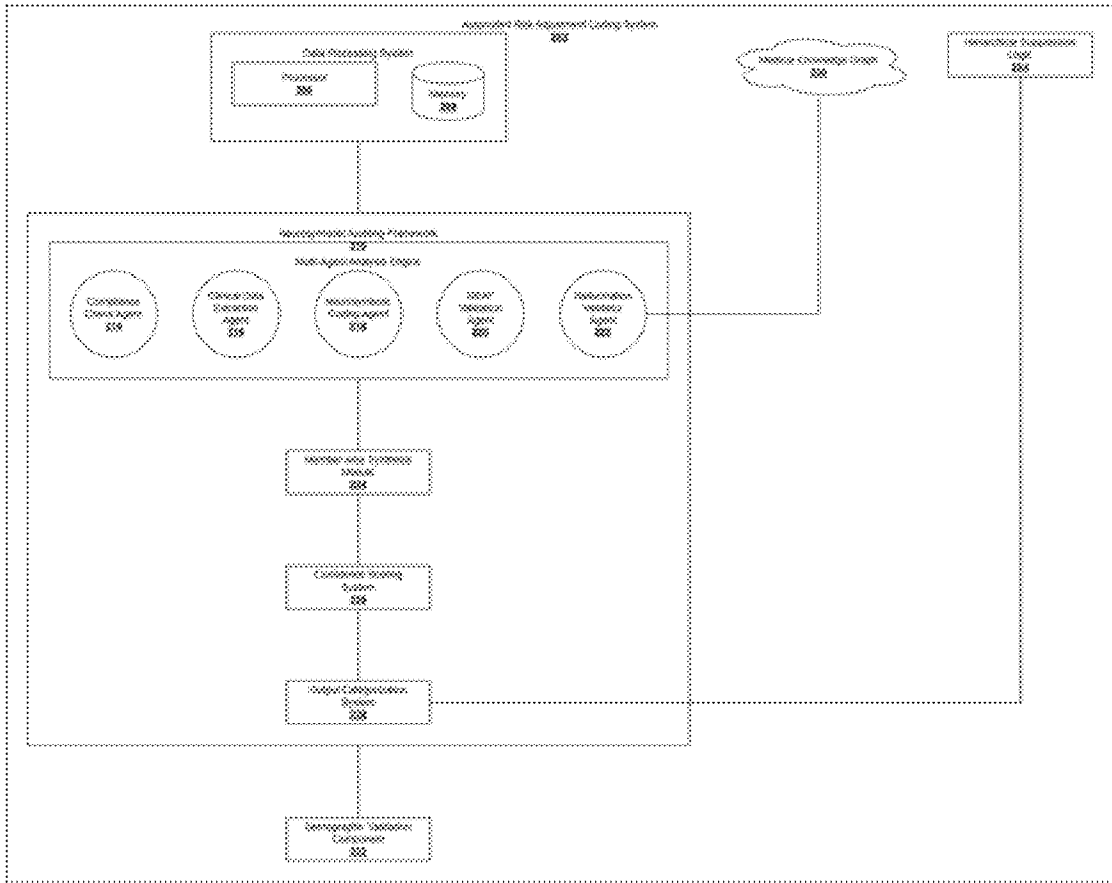


Figure 2