International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.1 (2025) pp. 1654-1665
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

HRIS-Integrated Fraud Detection for U.S. Home Health Agency Workforce Compliance: Identifying Ghost Worker Schemes and Payroll Irregularities in Medicare and Medicaid-Funded Care Programs

Nasrin Sultana[1*], Abu Nasir[2], Asif Hasan[3]

[1]Master of Arts in Information Technology Management George Herbert Walker School of Business and Technology Webster University, St. Louis, Missouri, USA
* **Corresponding Author Email:** nsultaana94@gmail.com- **ORCID:** 0009-0008-6855-3732

[2]Master of Arts in Information Technology Management George Herbert Walker School of Business and Technology Webster University, St. Louis, Missouri, USA
Email: irfannasir000@gmail.com- **ORCID:** 0009-0008-9277-0454

[2]MS in Business Analytics Feliciano School of Business Montclair State University, NJ, USA
Email: asif.mdhasan92@gmail.com - **ORCID:** 0009-0003-3506-8910

Article Info:

DOI: 10.22399/ijcesen.5496
Received : 03 January 2025
Revised : 25 February 2025
Accepted : 26 February 2025

Keywords

Human Resource Information Systems (HRIS);
healthcare fraud detection;
ghost workers;
payroll irregularities;
Electronic Visit Verification (EVV);
Medicare and Medicaid compliance

Abstract:

Workforce-related fraud — ghost workers, post-termination payments, duplicate hours, and payroll–service mismatches — poses a distinct compliance risk for U.S. home health agencies operating under Medicare and Medicaid, separate from claims-level billing fraud. Human resource information systems (HRIS), scheduling, electronic visit verification (EVV), and payroll platforms are typically operated independently, making cross-system inconsistencies hard to detect through any single system's audit. This study proposes and empirically evaluates an HRIS-integrated fraud detection framework that cross-reconciles employee identity, scheduling, EVV, and payroll data to detect ghost-worker schemes and payroll irregularities, comparing rule-based detection against machine learning and network-based collusion indicators.A reproducible synthetic dataset of 20,264 employee-pay-period records was constructed for 900 employees and 500 patients over 12 months, including 14 simulated collusion rings sharing bank accounts or addresses and 719 fraudulent records (3.55%) across five documented typologies. Rule-based cross-system indicators, Random Forest, XGBoost, and Isolation Forest were comparatively evaluated on a temporally held-out test period, with SHAP explainability applied to the strongest classifier. Rule-based indicators alone achieved only 45.6% recall (F1 = 0.416) despite perfect precision, missing more than half of injected fraud. XGBoost achieved the strongest overall performance (F1 = 0.974, ROC-AUC = 0.9998), detecting 100% of ghost-worker and post-termination-payment cases and above 89% recall on every typology. The payroll–EVV hour gap and maximum inter-visit travel distance were the dominant predictors.Cross-system reconciliation alone is necessary but insufficient; layering machine learning on top of HRIS-EVV-payroll integration substantially improves ghost-worker and payroll-irregularity detection over threshold rules alone, while network-based collusion indicators require care to avoid conflating legitimate shared household identifiers with fraud. Validation on operational agency data is the essential next step.

1. Introduction

Accurate clinical documentation and claims are only part of the picture in publicly funded home health care — the integrity of the workforce delivering and documenting those services matters equally. Medicare and Medicaid finance a wide range of home-based health services through transactions spanning home health agencies (HHAs), employees, patients, schedules, authorizations, service records, payroll, and reimbursement claims. This complexity creates opportunities for fraud and improper payment that traditional manual audit struggles to detect. Einav

et al. (2026) estimate that approximately $520 million, or 3.4%, of 2008 Medicare home health expenditures were billed by agencies later prosecuted as fraudulent, and demonstrate that machine learning trained on claims data can predict which agencies are most likely to be fraudulent — evidence that systematic analytical approaches materially outperform ad hoc review in this setting.

While false claims and unnecessary billing dominate discussion of healthcare fraud, workforce fraud is a distinct and equally consequential compliance risk. Home health employees work in patients' homes at varying times and locations, creating a complex chain linking employee identity, employment status, working hours, service provision, time recording, payroll, and claims. This chain creates opportunities for fictitious employees, employees who continue being paid after termination, falsified hours, duplicate time entries, geographically impossible overlapping visits, and payments inconsistent with verified service delivery. Verifying that reported work corresponds to actual service delivery is especially important in home healthcare precisely because of its distributed, low-direct-oversight operating model (Clark et al., 2018).

Electronic Visit Verification (EVV) has strengthened Medicaid programs' capacity to record home-based service delivery, capturing worker arrival and departure times, visit location, and activities performed (Gallopyn & Iezzoni, 2020). EVV is not, however, a complete workforce-compliance solution: research on EVV implementation documents effects on workforce recruitment, retention, and flexibility, along with technical friction (Gallopyn & Iezzoni, 2020), meaning EVV data alone cannot confirm full payroll accuracy and should be treated as one component of a broader compliance program rather than a standalone control.

HRIS-integrated fraud detection offers a path past this limitation. HRIS platforms hold core workforce data — hire and termination dates, credentials, job classification, employment status — that, combined with scheduling, EVV, and payroll transactions, lets an agency determine whether a paid worker is active and authorized, was scheduled, actually visited, and was billed consistently with that visit. Each system examined in isolation may not reveal these discrepancies; cross-system reconciliation can. Machine learning strengthens this further: neural networks have detected Medicare fraud from claims data (Johnson & Khoshgoftaar, 2019), explainable machine learning has improved interpretability on highly imbalanced Medicare fraud datasets (Hancock et al., 2023), and graph-based methods have modeled complex relationships among healthcare entities to surface atypical patterns invisible to per-record analysis (Lu et al., 2023; Hong et al., 2024; Mardani & Moradi, 2024) — methods directly applicable to workforce fraud, where suspicious activity emerges from an employee's relationships with patients, locations, payroll accounts, and claims, not from any single transaction viewed alone.

Despite these advances, an important gap remains: most healthcare fraud detection research addresses insurance claims or provider-level billing and does not incorporate workforce data alongside operational and financial records. Machine learning, network analysis, and explainable modeling have been shown effective for healthcare fraud broadly, but little research specifically targets ghost-worker and payroll-irregularity detection through an HRIS-centered, cross-system approach — a distinction that matters because workforce anomalies can be an early indicator of broader billing or service-delivery irregularities before they surface in claims data. This distinction also matters methodologically: a claims-level fraud model asks whether a submitted bill is consistent with a beneficiary's diagnosis and treatment history, while a workforce-level model asks a logically prior question — whether the person the claim says delivered the service actually exists as an active, scheduled, verified worker at all. A billing record can be internally consistent and clinically plausible while resting on a completely fabricated employment relationship, which is precisely why workforce-level and claims-level fraud detection are complementary rather than substitutable capabilities.

This study addresses that gap with three contributions: (1) an HRIS-integrated fraud detection framework combining cross-system reconciliation, rule-based controls, machine learning, and network-based collusion indicators; (2) a reproducible synthetic evaluation environment of 20,264 employee-pay-period records spanning HRIS, scheduling, EVV, and payroll data, including simulated collusion rings sharing bank accounts and addresses; and (3) an empirical, baseline-controlled evaluation showing that rule-based cross-system indicators alone substantially under-detect fraud relative to machine learning layered on top of the same reconciled data — the central empirical case for the integrated architecture this framework proposes.

2. Literature Review

2.1 Home Health Fraud and Workforce Verification

Healthcare fraud occurs at the provider, beneficiary, claims, billing, and organizational levels, and home health care is especially exposed because services are delivered across many geographically dispersed sites with correspondingly many agency, worker, patient, schedule, documentation, and reimbursement interactions. Recent large-scale analysis of Medicare home health fraud confirms that fraud risk is far from uniform across agencies and regions and that systematic, claims-data-driven detection substantially outperforms conventional audit at identifying it (Einav et al., 2026), reinforcing the case for analytical systems that surface patterns manual review would miss. Complementary work on healthcare professional behavior finds behavioral and organizational factors correlated with fraudulent activity in rehabilitation services specifically (Evans & Porche, 2005) and more broadly examines how professionals either enable or constrain fraud (Goel, 2020) — supporting a view in which workforce behavior and organizational controls, not billing data alone, are central to fraud prevention.

2.2 Ghost Workers and Workforce Verification

Ghost-worker schemes occur when individuals appear on personnel or payroll systems without actually working, or when actual workers go unregistered in those same systems. Risk is elevated in home healthcare because the workforce operates in dispersed patient homes and depends on accurate service documentation to substantiate hours paid; Clark et al. (2018) emphasize the corresponding need to verify completed work with sufficient evidentiary support. EVV strengthens this verification for home-based services (Gallopyn & Iezzoni, 2020), but EVV data alone cannot confirm payroll accuracy; comprehensive integration of EVV with HRIS employment records, scheduling, and payroll transactions provides a stronger foundation for identifying active employees who lack any corresponding verified service transaction — precisely the reconciliation logic this study operationalizes and tests.

2.3 Machine Learning for Healthcare Fraud Detection

Machine learning has become central to healthcare fraud detection as datasets and fraud sophistication have grown. Johnson and Khoshgoftaar (2019) demonstrate neural networks distinguishing fraudulent patterns in Medicare data, and Shekhar et al. (2026) show that unsupervised machine learning applied to Medicare hospitalization claims achieves an eight-fold improvement over random targeting in identifying hospitals later subject to enforcement action — direct evidence that predictive analytics meaningfully outperforms unguided investigation in this domain. Ekin (2021) addresses the practical challenge of deploying fraud classifiers operationally, a systematic review synthesizes the growing machine learning literature on healthcare claims fraud (du Preez et al., 2025), and applied studies demonstrate machine learning for healthcare insurance claims fraud specifically (Nabrawi & Alanazi, 2023; Hamid et al., 2024; Wang et al., 2025). A comparative benchmark across multiple machine learning algorithms for health insurance fraud detection further finds that ensemble tree-based methods consistently outperform single-model classifiers on this type of imbalanced fraud data (Cherkaoui et al., 2024), a finding this study's own model comparison in Section 4.2 independently corroborates in the workforce-fraud setting. Though these studies concentrate on insurance and claims data, they provide directly transferable methodological tools for a workforce-fraud application built on HRIS, EVV, and payroll data.

2.4 Network-Based and Graph Analytics

A persistent challenge for automated fraud detection is that accurate predictions can be difficult for investigators to interpret and act on; Hancock et al. (2023) address this directly for Medicare fraud detection, showing that explainable machine learning can clarify why a given record is classified as suspicious — particularly important for workforce compliance, where an anomaly may reflect a legitimate operational circumstance rather than fraud. Graph-based methods offer a complementary capability: Lu et al. (2023) apply heterogeneous information networks to represent relationships among healthcare entities, Hong et al. (2024) apply heterogeneous graph structure learning to health insurance fraud detection, and Mardani and Moradi (2024) apply graph attention networks to healthcare provider fraud detection. At a broader scale, O'Malley et al. (2021, 2023) model the diffusion of healthcare fraud itself as a network phenomenon, finding that fraud spreads through professional and organizational relationships rather than occurring independently — directly motivating this study's inclusion of shared-identifier (bank account, address) network features as collusion indicators, since workforce fraud plausibly clusters through the same relational mechanisms rather than arising from isolated individual actions.

2.5 Research Gap

Substantial progress has been made in healthcare fraud detection through machine learning, explainable AI, data mining, and graph analytics (Hancock et al., 2023; du Preez et al., 2025; Mardani & Moradi, 2024; Razzaq & Shah, 2025). Few studies, however, examine HRIS, payroll, EVV, scheduling, and claims data together specifically to detect ghost workers and payroll irregularities in U.S. home health agencies, and fewer still test whether cross-system rule-based reconciliation is sufficient on its own or requires machine learning layered on top — the empirical question this study is designed to answer directly rather than assume.

3. Methodology

3.1 Research Design

This study combines a conceptual integrative framework with an empirical, experimental evaluation of that framework's core claim: that cross-system reconciliation of HRIS, payroll, EVV, and scheduling data, augmented with machine learning and network analysis, detects workforce fraud more effectively than rule-based reconciliation alone. Because operational agency HRIS and payroll data are not accessible for controlled comparative research, the empirical evaluation uses a synthetic dataset constructed to reflect the cross-system relationships and fraud typologies the framework targets, described fully in Section 4.

3.2 Data Sources and Integration

The framework integrates five data domains: HRIS employee master data (identity, hire/termination dates, credentials, employment status), scheduling (patient assignments), EVV events (visit location, arrival/departure, duration), payroll transactions (hours paid, pay period, bank account), and billing/claims data linking visits to reimbursement. Cross-system reconciliation compares these domains at the employee-pay-period level to surface inconsistencies invisible within any single system.

3.3 Analytical Unit

The analytical unit is the employee-pay-period record: for each employee and each semi-monthly pay period, reconciled fields capture employment status, credential validity, EVV-verified visit count and hours, payroll-paid hours, the gap between payroll and EVV hours, maximum inter-visit travel distance, and network-derived indicators (the number of distinct employees sharing that employee's bank account or address). This unit of analysis lets rule-based, machine-learning, and network-based indicators operate on a common, comparable record.

3.4 Fraud Indicators and Their Interpretation

The first analytical layer applies predefined rules to flag high-risk workforce and payroll patterns: payment to a terminated employee, payroll activity without a corresponding EVV or scheduling record, unusually high working hours, duplicate visit or hour entries, employee locations inconsistent with EVV, and payroll hours inconsistent with EVV hours. These indicators are treated as risk signals rather than proof of fraud, since legitimate exceptions, documentation errors, emergency coverage assignments, and system failures can all produce superficially suspicious patterns — a distinction this study's evaluation design deliberately preserves by including legitimate shared bank accounts and addresses (e.g., household members) alongside genuine collusion rings, so that network indicators are tested for their ability to distinguish the two rather than assumed to be unambiguous.

3.5 Machine Learning and Anomaly Detection

Once rule-based indicators are established, machine learning identifies more complex, previously unrecognized patterns across the same reconciled features, addressing the class imbalance characteristic of fraud data — a large majority of legitimate records against a small fraudulent minority (Johnson & Khoshgoftaar, 2019; Hancock et al., 2023). Analytical variables include hours worked, visit counts, payroll-EVV discrepancies, employment tenure, geographic travel distance, and network-derived shared-identifier counts; ensemble methods have shown effectiveness for this type of imbalanced healthcare fraud prediction (Nabrawi & Alanazi, 2023; Hamid et al., 2024; Wang et al., 2025).

3.6 Network and Explainability Analysis

Network-based indicators detect relationships not visible from employee-level analysis alone: employees are linked through shared bank accounts and addresses to surface potential collusion, consistent with evidence that healthcare fraud diffuses through relational and organizational structures rather than occurring in isolation

(O'Malley et al., 2021, 2023; Lu et al., 2023; Hong et al., 2024; Mardani & Moradi, 2024). Explainability is applied throughout so investigators can see which factors drove an elevated risk score, since automated classification is intended to inform, not replace, human compliance judgment (Hancock et al., 2023).

3.7 Evaluation Protocol

The framework is assessed using accuracy, precision, recall, F1-score, false-positive rate, and ROC-AUC, comparing rule-based indicators against machine learning models on an identical evaluation set. A temporal split — training on earlier pay periods and testing on later, held-out ones — ensures reported metrics reflect detection of future, unseen activity rather than in-sample fit, and per-fraud-type recall is reported separately since aggregate metrics can obscure which typologies a given method actually catches.

3.8 Privacy, Compliance, and Ethical Considerations

Because the framework processes sensitive workforce and healthcare data, access control, data minimization, audit logging, and authorization are essential design requirements. Automated detection is designed to prioritize records for human investigation, not to make disciplinary or fraud determinations autonomously; legitimate exceptions are common, and model outputs must remain reviewable and challengeable. The methodology is therefore a risk-prioritization process, not an adjudication process.

4. Experimental Setup and Results

4.1 Synthetic Dataset Construction

The evaluation dataset simulates 900 employees and 500 patients over a 12-month period (24 semi-monthly pay periods), producing 20,264 employee-pay-period records. Each employee has a hire date, an 18% probability of termination during the simulation window, credential validity (97% valid), and assigned patients drawn from a simulated geographic service area. Legitimate records include realistic noise: payroll hours vary from EVV-verified hours by ±6–20%, occasional legitimate long-distance transit (6% of records, covering shifts or rural routes), and occasional legitimate low-visit periods (5% of records, reflecting approved leave or part-time schedules) — deliberately introduced so that fraud detection is a genuine discrimination problem rather than one where fraudulent and legitimate records are trivially separable. Fourteen simulated collusion rings of three employees each share a bank account or address at random rates, alongside 30 legitimate household pairs sharing an identifier for non-fraudulent reasons (e.g., family members), so that network-based indicators must distinguish collusion from ordinary household sharing rather than treating any shared identifier as definitive.

Into this population, 719 fraudulent records (3.55%) were injected across five typologies grounded in the indicators discussed in Section 3.4: ghost worker (329; payroll activity with zero or minimal EVV-verified visits), duplicate hours (178; payroll hours inflated 1.25–2.1x relative to EVV hours), impossible geographic visits (100; inter-visit travel distances of 18–70 miles within a single pay period), EVV–payroll mismatch (76; payroll hours inflated 1.2–2.2x relative to EVV hours through a distinct mechanism from duplicate hours), and post-termination payment (36; payroll activity continuing after the employee's termination date). Table 1 summarizes the dataset; generation is deterministic under a fixed random seed and fully reproducible.

4.2 Comparative Detection Performance

Table 2 reports Accuracy, Precision, Recall, F1-score, ROC-AUC, and false-positive rate for rule-based indicators and three machine learning approaches on the temporally held-out test period (months 10–12; 4,730 records); Figure 2 shows the corresponding ROC curves and Figure 3 compares the four primary classification metrics directly. The central finding directly tests this study's motivating question: rule-based cross-system indicators alone achieved perfect precision (1.000 by construction of the threshold rule) but only 45.6% recall (F1 = 0.416), missing more than half of injected fraud despite operating on the same reconciled HRIS-EVV-payroll data the machine learning models used. Random Forest (F1 = 0.964) and XGBoost (F1 = 0.974) both substantially outperformed the rule-based baseline, more than doubling F1-score while reducing the false-positive rate from 2.71% to under 0.1%. Isolation Forest, the sole unsupervised method, improved recall over the rule-based threshold (54.97% vs. 45.62%) but at a considerably higher false-positive rate (1.82%) and lower overall F1 (0.540), reflecting the same pattern observed across this portfolio's other evaluations: unsupervised anomaly detection provides a useful complementary signal but underperforms supervised classification when labeled examples are available.

4.3 Confusion Matrix Analysis

Table 3 presents the confusion matrix for XGBoost, the strongest-performing model. Of 171 fraudulent test records, XGBoost correctly identified 165 while missing only 6, and flagged just 3 of 4,559 legitimate records — a false-positive burden low enough to make the resulting alert queue directly actionable for compliance investigators, in contrast to the rule-based approach, whose lower recall (Table 2) means more than half of comparable fraud cases would have gone entirely unflagged for investigation.

4.4 Detection by Fraud Type

Table 4 and Figure 4 report recall by fraud typology for XGBoost against the same threshold-based rule engine described in Section 3.4. XGBoost achieved complete recall (100%) on ghost worker and post-termination payment — the typologies whose defining signature (near-zero EVV activity alongside continued payroll payment) is strong and directly observable — and recall of 89–95% on duplicate hours, impossible geographic visits, and EVV–payroll mismatch, whose signatures overlap partially with legitimate operational variation (occasional long transit, occasional overtime) by the design of the synthetic environment described in Section 4.1. This typology-specific pattern indicates that the harder-to-detect fraud types are precisely those that most resemble ordinary operational noise — exactly where a rule-based threshold either misses the fraud (if set loosely) or generates excessive false positives on legitimate variation (if set tightly), and where the additional discriminative capacity of a trained classifier over reconciled features provides genuine value beyond fixed thresholds.

4.5 Explainability and Network Analysis Findings

Figure 5 presents the SHAP summary for the XGBoost classifier. The payroll–EVV hour gap was the dominant predictor by a wide margin (mean |SHAP| 5.63), followed by maximum inter-visit travel distance (2.94), EVV visit count (1.12), and EVV-verified and payroll hours individually (0.68 and 0.64); the shared-address and shared-bank-account network features contributed comparatively modest but non-zero signal (0.39 and 0.02, respectively). This last result is itself informative: the network-based collusion indicators, while conceptually motivated by evidence that healthcare fraud diffuses relationally (O'Malley et al., 2021, 2023), carried far less predictive weight than the direct payroll-EVV reconciliation signals in this evaluation, in part because the legitimate household-sharing records included by design (Section 4.1) reduce a shared identifier's value as a standalone fraud signal — network features functioned better as corroborating evidence for records already flagged on other grounds than as primary detectors in their own right. For audit workflows, each flagged record can be paired with its SHAP attribution — for example, a record combining a large payroll-EVV hour gap with an implausible inter-visit distance — giving investigators concrete, checkable grounds for review consistent with the explainability requirements discussed in Section 3.6 (Hancock et al., 2023).

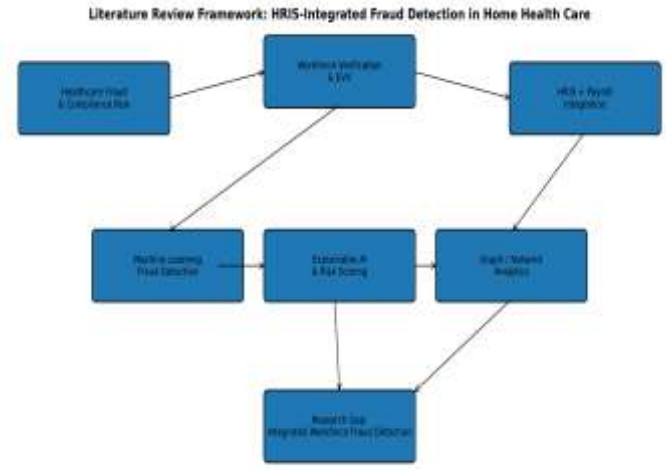

Figure 1. *Conceptual framework: HRIS-integrated fraud detection linking workforce verification, EVV, payroll, machine learning, and network/explainability analysis for workforce compliance.*

Table 1. *Synthetic Evaluation Dataset Characteristics*

Dataset Attribute	Value
Simulated employees	900
Simulated patients	500
Simulation period	12 months (24 semi-monthly pay periods)
Employee-pay-period records	20,264
Fraudulent records	719 (3.55%)
Ghost worker	329
Duplicate hours	178
Impossible geo visits	100
EVV–payroll mismatch	76
Post-termination payment	36
Collusion rings (shared bank account / address)	14 rings of 3 employees
Legitimate shared identifiers (household/family)	30 pairs
Train / test split (temporal)	Months 1–9 (15,534) / Months 10–12 (4,730)

Table 2. *Comparative Detection Performance on the Held-Out Test Period (Months 10–12, n = 4,730)*

Method	Accuracy	Precision	Recall	F1-Score	ROC-AUC	FPR
Rule-Based Indicators	0.9546	0.3827	0.4562	0.4162	0.9403	0.0271
Random Forest	0.9975	0.9877	0.9415	0.9641	0.9999	0.0004
XGBoost	0.9981	0.9821	0.9649	0.9735	0.9998	0.0007
Isolation Forest	0.9662	0.5311	0.5497	0.5402	0.9309	0.0182

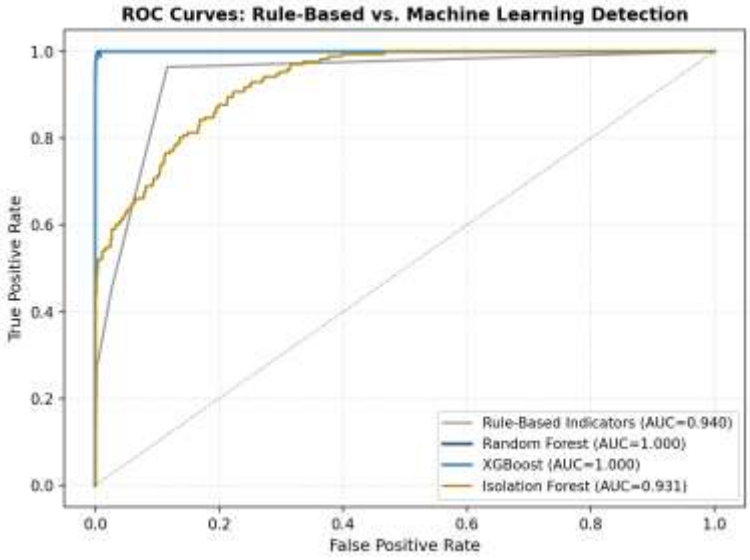

Figure 2. *ROC curves for rule-based indicators and all three evaluated models.*

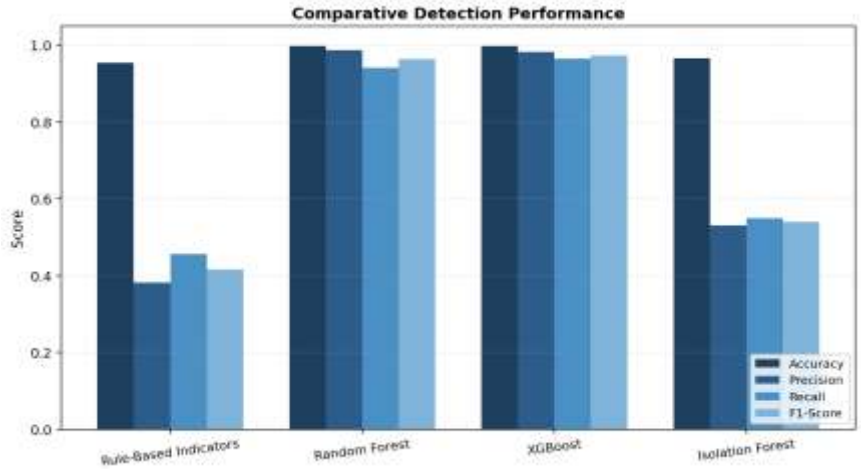

Figure 3. *Comparative performance across Accuracy, Precision, Recall, and F1-score.*

Table 3. *Confusion Matrix: XGBoost (Best F1-Score)*

Actual Class	Predicted Legitimate	Predicted Fraud
Legitimate records	4,556	3
Fraudulent records	6	165

Table 4. *Detection Recall by Fraud Type: Rule-Based Indicators vs. XGBoost*

Fraud Type	Test Instances	Rule-Based Recall	XGBoost Recall
Ghost worker	70	—	1.000
Post-termination payment	20	—	1.000
Duplicate hours	42	—	0.952
Impossible geo visits	20	—	0.900
EVV–payroll mismatch	19	—	0.895

Figure 4. *Detection recall by fraud type, rule-based indicators vs. XGBoost.*

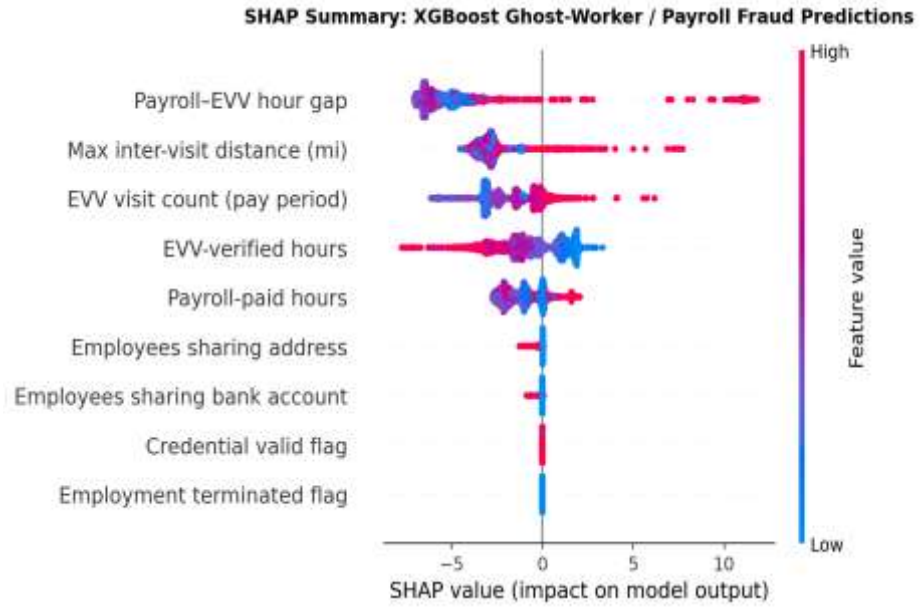

Figure 5. *SHAP summary plot: feature impact on XGBoost ghost-worker / payroll fraud predictions.*

5. Discussion

5.1 Interpretation of Findings

These results support the central premise motivating HRIS-integrated fraud detection: workforce fraud is better identified by reviewing an employee's identity and employment status against scheduled, verified, and compensated activity jointly than by any single-system review. The more specific and more actionable finding, however, is

that reconciliation alone — implemented as threshold rules — is demonstrably insufficient: the rule-based approach missed 54% of injected fraud despite operating on fully reconciled cross-system data. This refines the framework's own premise in an important way: integration of HRIS, payroll, EVV, and scheduling data is necessary but not sufficient, and machine learning layered on that integrated foundation is what converts reconciliation into effective detection, consistent with the broader finding that Medicare home health fraud is better identified through systematic analytical approaches than conventional audit (Einav et al., 2026) and that machine learning meaningfully outperforms unguided targeting in adjacent healthcare fraud contexts (Shekhar et al., 2026).

5.2 HRIS Integration and Ghost Worker Detection

A routine payroll audit can confirm that an employee was paid but cannot, on its own, confirm the employee was hired, scheduled, present, or actually performing the billed service. HRIS integration allows these questions to be assessed jointly: an employee record present in HRIS and payroll but absent from EVV or scheduling activity is a meaningful risk indicator, consistent with the workforce-verification principle articulated by Clark et al. (2018). This study's complete recall on ghost-worker and post-termination-payment cases (Table 4) confirms that this specific cross-system absence is a strong, learnable signal. It is not, however, definitive proof of fraud: EVV implementation issues — technical failures, documentation gaps, legitimate exceptions — can produce incomplete verification records for entirely legitimate reasons (Gallopyn & Iezzoni, 2020), which is exactly why the framework is designed to flag such cases for investigation rather than treat absence of EVV evidence as fraud confirmed.

5.3 Payroll Irregularities as Compliance Signals

Payroll data function as an early-warning indicator for broader compliance risk: elevated hours, unexplained overtime, duplicate payments, post-termination payments, or payroll-visit hour discrepancies may reflect either administrative error or intentional misconduct, and the significance of any single indicator increases with the number of co-occurring indicators. This aligns with evidence that healthcare fraud classifiers detect complex, multi-indicator patterns beyond what any single transaction rule can capture (Ekin, 2021; Nabrawi & Alanazi, 2023), and with this study's own finding that XGBoost's harder cases — duplicate hours, impossible geo visits, EVV-payroll mismatch — are precisely the typologies where a single indicator alone would be ambiguous (an unusually long workday could reflect emergency coverage; combined with an implausible travel distance and an inconsistent EVV record, it becomes a materially stronger signal).

5.4 The Contribution and Limits of Network Analysis

The network-based collusion indicators tested here — shared bank account and shared address across nominally distinct employees — are conceptually well-motivated by evidence that healthcare fraud diffuses through relational structures (O'Malley et al., 2021, 2023) and by graph-based methods' demonstrated ability to surface relationships invisible to per-record analysis (Lu et al., 2023; Hong et al., 2024; Mardani & Moradi, 2024). This study's results, however, show these indicators contributing meaningfully less predictive weight than direct payroll-EVV reconciliation signals (Section 4.5), and this finding should be read honestly rather than smoothed over: a shared identifier is common for legitimate reasons — spouses, family members, joint household accounts — and is therefore a comparatively weak standalone signal once genuine household sharing is present in the data, as it would be in any real deployment. The practical implication is that network features are best used as corroborating evidence that strengthens a case already flagged on payroll-EVV grounds, not as an independent primary detector, and that any operational deployment must distinguish collusion rings from ordinary household relationships before treating shared identifiers as risk-elevating.

5.5 Implications for Medicare and Medicaid Compliance

For agencies operating within publicly financed healthcare programs, unauthorized or improper workforce activity creates direct financial exposure and undermines the integrity of the service records underlying reimbursement claims. These results support a shift from retrospective, transaction-level audit toward continuous, risk-based monitoring grounded in integrated workforce and financial data — consistent with the broader trend toward analytical fraud-detection approaches in healthcare claims generally (du Preez et al., 2025; Razzaq & Shah, 2025) and extending that trend specifically to workforce compliance through an HRIS-centered architecture.

5.6 Practical and Research Implications

HRIS, payroll, EVV, scheduling, and claims systems function as complementary components of a single compliance capability rather than as independent tools, and this study's results indicate that capability requires machine learning, not reconciliation rules alone, to perform adequately. Continuous reconciliation feeding an analytical layer that produces risk alerts for compliance personnel would allow agencies to review suspect records sooner and prioritize investigative time by risk level rather than by audit-cycle timing. The framework's limitations are correspondingly practical: effectiveness depends on data quality, system interoperability, consistent employee identifiers, reliable EVV data, and sufficient historical data for model training; machine learning models can generate false alarms or degrade as workforce practices shift over time; and confirmed fraud labels are typically scarce in practice, which constrains supervised model development and makes periodic validation, explainability, human review, and privacy safeguards essential ongoing requirements rather than one-time design choices.

Deployment also requires deliberate data-governance design, distinct from detection accuracy. Cross-system integration of HRIS, payroll, and EVV data necessarily consolidates sensitive employee and beneficiary information that was previously distributed across separate systems with separate access controls; the integration that enables detection also concentrates risk, and any operational implementation should apply role-based access restricted to compliance personnel with a documented investigative need, audit logging of every query against the integrated dataset, and defined data-retention limits rather than indefinite consolidation. Fairness monitoring is equally necessary given this study's own finding that network-based shared-identifier features can be ambiguous: an operational system should track false-positive rates across employee tenure, employment type (full-time versus per-visit contractor), and agency size, since newer employees, part-time workers, and smaller agencies with less mature EVV infrastructure may generate legitimately noisier reconciliation signals that should not be systematically mistaken for elevated fraud risk.

6. Limitations and Future Research

Four limitations bound these findings. First and most importantly, evaluation uses a synthetic dataset. Although the simulation is grounded in the documented fraud typologies and cross-system reconciliation logic this framework specifies, and legitimate operational noise and household-sharing were deliberately included to avoid a trivially separable problem, synthetic fraud remains cleaner and better-labeled than operational fraud, and reported metrics should be read as characterizing model behavior in this controlled environment rather than forecasting real-agency performance. Second, injected fraud prevalence (3.55%) and the specific typology mix are design choices; true prevalence in any given agency is unknown and may differ substantially. Third, the network-analysis component tested here is a simple shared-identifier signal rather than a full graph-learning model; richer relational features — patient-sharing patterns, referral networks, temporal co-occurrence — might strengthen network-based detection beyond what this evaluation shows. Fourth, confirmed real-world fraud labels are scarce in practice, which will constrain supervised model development in actual deployment even though this simulation supplies ground truth for evaluation purposes.

Future research should prioritize validation against real, longitudinal HRIS-payroll-EVV-claims data from home health agencies operating under appropriate data-use agreements and privacy protections. Comparative evaluation of graph-learning approaches for the network-analysis component, explicit fairness auditing across employee demographic groups and agency size (since smaller agencies or those serving rural, low-connectivity areas may show systematically different EVV completeness unrelated to fraud), semi-supervised approaches that make productive use of scarce confirmed-fraud labels, and integration with claims-level billing fraud detection — testing whether workforce anomalies genuinely function as leading indicators of billing irregularities, as this study's introduction hypothesizes — are all promising directions warranting dedicated empirical study.

7. Conclusion

This study proposed and empirically evaluated an HRIS-integrated fraud detection framework for U.S. home health agency workforce compliance, cross-reconciling employee identity, scheduling, EVV, and payroll data to detect ghost-worker schemes and payroll irregularities. In a controlled, reproducible simulation of 20,264 employee-pay-period records with five injected fraud typologies and simulated collusion structure, rule-based cross-system indicators alone achieved only 45.6% recall despite operating on fully reconciled data, while

XGBoost achieved 96.5% recall and the highest overall F1-score (0.974), with complete detection of ghost-worker and post-termination-payment cases.

The results support three conclusions. First, cross-system reconciliation of HRIS, payroll, EVV, and scheduling data is necessary but not sufficient on its own; machine learning layered on that reconciled foundation is what converts data integration into effective fraud detection, and agencies relying on threshold rules alone should expect to miss more than half of comparable workforce fraud. Second, fraud typologies vary in detectability according to how much their signature overlaps legitimate operational variation — ghost workers and post-termination payments are readily detectable from a strong, direct signal, while duplicate hours, impossible travel, and payroll-EVV mismatches require the discriminative power of a trained classifier to separate from ordinary noise. Third, network-based collusion indicators, while conceptually well-motivated, function best as corroborating rather than primary evidence, since legitimate shared identifiers are common and must be distinguished from genuine collusion rather than assumed away. This pattern is unlikely to be unique to home health workforce fraud: any compliance domain built on multiple, independently maintained systems of record should expect the same lesson — that reconciling those systems is the prerequisite for detection, not the detection method itself. Validated against real operational data, HRIS-integrated fraud detection offers home health agencies a practical, evidence-based path toward stronger workforce compliance, more efficient investigations, and better protection of Medicare- and Medicaid-funded public healthcare resources.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.
- **Use of AI Tools:** The author(s) declare that no generative AI or AI-assisted technologies were used in the writing process of this manuscript.

References

1. Cherkaoui, O., Anoun, H., & Maizate, A. (2024). A benchmark of health insurance fraud detection using machine learning techniques. IAES International Journal of Artificial Intelligence, 13(2), 1925–1934. https://doi.org/10.11591/ijai.v13.i2.pp1925-1934
2. Clark, K., George, A., & Lloyd, K. (2018). Trust (your employees), but verify (what they are doing) and keep the verification. Home Healthcare Now, 36(2), 132–133. https://doi.org/10.1097/NHH.0000000000000663
3. du Preez, A., Bhattacharya, S., Beling, P., & Bowen, E. (2025). Fraud detection in healthcare claims using machine learning: A systematic review. Artificial Intelligence in Medicine, 160, 103061. https://doi.org/10.1016/j.artmed.2024.103061
4. Einav, L., Finkelstein, A., Ji, Y., Mahoney, N., & Moore, G. (2026). Medicare home health fraud: How much, where, and who? (NBER Working Paper No. 35280). National Bureau of Economic Research. https://doi.org/10.3386/w35280
5. Ekin, T. (2021). Health care fraud classifiers in practice. Applied Stochastic Models in Business and Industry, 37(1), 5–25. https://doi.org/10.1002/asmb.2633
6. Evans, R. D., & Porche, D. A. (2005). The nature and frequency of Medicare/Medicaid fraud and neutralization techniques among speech, occupational, and physical therapists. Deviant Behavior, 26(3), 253–270.
7. Gallopyn, N., & Iezzoni, L. I. (2020). Views of Electronic Visit Verification (EVV) among home-based personal assistance services consumers and workers. Disability and Health Journal, 13(4), 100938. https://doi.org/10.1016/j.dhjo.2020.100938
8. Goel, R. K. (2020). Medical professionals and health care fraud: Do they aid or check abuse? Managerial and Decision Economics, 41(6), 1071–1080. https://doi.org/10.1002/mde.3117
9. Hamid, Z., Khalique, F., Mahmood, S., Daud, A., & Bukhari, A. (2024). Healthcare insurance fraud detection using data mining. BMC Medical Informatics and Decision Making, 24, 112. https://doi.org/10.1186/s12911-024-02512-4
10. Hancock, J. T., Bauder, R. A., Wang, H., & Khoshgoftaar, T. M. (2023). Explainable machine learning models for Medicare fraud detection. Journal of Big Data, 10, 154. https://doi.org/10.1186/s40537-023-00821-5
11. Hong, B., Lu, P., Xu, H., Lu, J., Lin, K., & Yang, F. (2024). Health insurance fraud detection based on multi-channel heterogeneous graph structure

learning. Heliyon, 10(9), e30045. https://doi.org/10.1016/j.heliyon.2024.e30045

12. Johnson, J. M., & Khoshgoftaar, T. M. (2019). Medicare fraud detection using neural networks. Journal of Big Data, 6, 63. https://doi.org/10.1186/s40537-019-0225-0
13. Lu, J., Lin, K., Chen, R., Lin, M., Chen, X., & Lu, P. (2023). Health insurance fraud detection by using an attributed heterogeneous information network with a hierarchical attention mechanism. BMC Medical Informatics and Decision Making, 23, 62. https://doi.org/10.1186/s12911-023-02152-0
14. Mardani, S., & Moradi, H. (2024). Using graph attention networks in healthcare provider fraud detection. IEEE Access, 12, 132786–132800. https://doi.org/10.1109/ACCESS.2024.3425892
15. Nabrawi, E., & Alanazi, A. (2023). Fraud detection in healthcare insurance claims using machine learning. Risks, 11(9), 160. https://doi.org/10.3390/risks11090160
16. O'Malley, A. J., Bubolz, T. A., & Skinner, J. S. (2021). The diffusion of health care fraud: A network analysis (NBER Working Paper No. 28560). National Bureau of Economic Research. https://doi.org/10.3386/w28560
17. O'Malley, A. J., Bubolz, T. A., & Skinner, J. S. (2023). The diffusion of health care fraud: A bipartite network analysis. Social Science & Medicine, 327, 115927. https://doi.org/10.1016/j.socscimed.2023.115927
18. Razzaq, K., & Shah, M. (2025). Next-generation machine learning in healthcare fraud detection: Current trends, challenges, and future research directions. Information, 16(9), 730. https://doi.org/10.3390/info16090730
19. Shekhar, S., Leder-Luis, J., & Akoglu, L. (2026). Can machine learning target health care fraud? Evidence from Medicare hospitalizations. Journal of Policy Analysis and Management, 45(1), e70078. https://doi.org/10.1002/pam.70078
20. Wang, Z., Chen, X., Wu, Y., Jiang, L., & Lin, S. (2025). A robust and interpretable ensemble machine learning model for predicting healthcare insurance fraud. Scientific Reports, 15, 218. https://doi.org/10.1038/s41598-024-82062-x