HMPI | HEALTH MANAGEMENT, POLICY & INNOVATION

BAHM BUSINESS SCHOOL ALLIANCE FOR HEALTH MANAGEMENT

The Emerging Risk of AI-Enabled Healthcare Fraud

Elman Amador Medina, Graduate School of Business, Stanford University; Steven Gong, Dashiell Miner, School of Medicine, Stanford University; Lauren Bilbo, Klara Klarowicz, Graduate School of Business, Stanford University; Sergio Mavridis, David Scheinker, Stefanos Zenios, Graduate School of Business, Stanford University; Kevin Schulman, Graduate School of Business, School of Medicine, Stanford University

Support: The Ludy Family Foundation, The Hirsch Family Foundation, The Mindshare Institute, and the Government, Business and Society Initiative of the Graduate School of Business, Stanford University.

Abstract

What is the message? Healthcare fraud has always been a major concern for health payers. While public debate often equates healthcare fraud with improper beneficiary enrollment, most documented losses arise from provider-side and organized criminal schemes. The structural complexity of the U.S. healthcare and lack of governance over the market has resulted in significant lags in response emerging artificial intelligence (AI)-enabled fraud. This paper proposes a technical framework to transition the industry from reactive fraud detection to proactive fraud prevention by adapting multi-layered defense mechanisms proven effective in the financial services sector.

What is the evidence? The proposed solution comprises a three-layer architecture integrated into claims processing: (1) Healthcare Identity Management, analogous to Know Your Customer (KYC) protocols, for rigorous provider and patient identity verification; (2) Continuous Claim Validation using “Zero Trust” principles, where each claim is assumed to be fraudulent until proven valid; and (3) AI/Machine Learning (ML)-Powered Real-Time Detection using advanced analytics and collaborative intelligence to identify aberrant patterns before payment authorization.

HMPI | HEALTH MANAGEMENT, POLICY & INNOVATION
BAHM BUSINESS SCHOOL ALLIANCE FOR HEALTH MANAGEMENT

Timeline: May 31, 2026; accepted after review June 12, 2026.

Cite as: Elman Amador Medina, Steven Gong, Dashiell Miner, Lauren Bilbo, Klara Klarowicz, Sergio Mavridis, David Scheinker, Stefanos Zenios, Kevin Schulman. 2026. The Emerging Risk of AI-Enabled Healthcare Fraud. Health Management, Policy and Innovation (www.HMPI.org). Volume 11, Issue 2.

Introduction

Transaction fraud in claims and reimbursement is an unfortunate feature of the U.S. healthcare market. Improper payments, generally attributed to fraud and abuse, are thought to exceed $100 billion annually across the Medicare and Medicaid programs [1]. Fraud is criminal activity to steal money through the insurance claims process, while abuse describes non-criminal clinical practices or billing patterns. This paper focuses on the issue of payment fraud.

Healthcare fraud takes two distinct forms. Provider-side fraud occurs when clinicians, suppliers, or organized criminal groups bill payers. Beneficiary fraud occurs when individuals enroll in coverage or use services to which they are not entitled.

While recent policy debates, including the Medicaid eligibility provisions of the One Big Beautiful Bill Act of 2025, have centered on beneficiary fraud, enforcement data tell a different story [2]. The large majority of documented fraud losses come from provider-side schemes and organized criminal enterprises [3-5]. Several patterns of provider-side fraud have been identified: medical identity theft, billing for unnecessary services, billing for services not provided, upcoding, unbundling, and kickbacks [3]. Historically, fraud has been a particular issue in the Medicare program, where Medicare must pay for services within a fixed period of time. When it discovers that it paid fraudulent claims, it must “chase” after the biller for a recovery (the so-called pay and chase model). Medicare began to check for fraud prior to payment in 2011 [6], but fraudulent billing still persists [4].

AI-Enabled Fraud

These legacy issues of fraudulent billing are now facing new challenges with the emergence of artificial intelligence (AI). The convergence of generative AI and healthcare's analog infrastructure is creating an unprecedented fraud escalation.

According to data security vendors, deepfake fraud[1] incidents surged globally by over 700% between early 2024 and early 2025, with North America experiencing a 1,100% increase over the same period [7]. Synthetic identity document fraud—in which AI-generated credentials combine real and fabricated personal data—rose 311% in North America during Q1 2025 compared to Q1 2024 [7].

These capabilities are already being weaponized against healthcare payment systems. Deepfake medical identities enable phantom billing schemes, fraudulent prescription access, and insurance claims filed under fabricated patient profiles. Voice cloning technology, which now requires as little as three to five seconds of audio to produce a convincing replica, has driven a 475% year-over-year increase in attacks targeting insurers [8]. Telemedicine platforms present particular vulnerabilities; the 2025 DOJ National Health Care Fraud Takedown included $1.17 billion in fraudulent telemedicine and genetic testing claims, with foreign actors using AI to generate fraudulent audio recordings of Medicare beneficiary consent [4]. In August 2025, the Department of Justice announced its first healthcare AI fraud enforcement action against Troy Health, Inc., which used an AI platform (Troy.ai) to fraudulently enroll over 300 Medicare beneficiaries in a single day without their knowledge or consent—a harbinger of how AI reduces the marginal cost of fraud execution [9]. CMS prevented over $4 billion in fraudulent claims from being paid in the lead up to the June 2025 enforcement action [10].

Nation State Threat Actors

Healthcare's vulnerability extends beyond financially motivated criminal enterprises to include advanced persistent threat (APT) groups operating with the tacit or explicit support of foreign governments. These groups can target intellectual property for theft and commit financial crimes.

Iranian state-sponsored group Pioneer Kitten provides network access to ransomware affiliates who deploy attacks against healthcare systems, creating a cooperative model between nation-

state espionage and criminal extortion [11]. North Korean actors deployed the Maui ransomware specifically against U.S. hospitals between 2021 and 2022, with ransom payments funding the state's nuclear weapons program [12]. Chinese APT groups have targeted biomedical research institutions for COVID-19 vaccine intellectual property theft and deployed malware through medical imaging DICOM files [13]. Russian-linked APT29 (Cozy Bear), attributed to the Russian Foreign Intelligence Service (SVR), has conducted persistent espionage campaigns against healthcare and pharmaceutical organizations in the United States and Europe [14]. Ransomware attacks against healthcare organizations have increased 300% since 2015, and in 2024, healthcare reported the most cyber threat incidents of any sector: 458 ransomware incidents targeting the health sector [15][16].

In June 2026, Google sued a Chinese enterprise, Outsider Enterprise, for using Gemini's AI tools to create fake websites including websites that appear like government sites including the Postal Service and New York State's E-Z Pass system. Google reported that hundreds of thousands of people had been exposed to these fraudulent sites. Google was cooperating with the Federal Bureau of Investigation in the case, with wireless carriers to try to shut down the network. [17]

These APT groups could use AI technology to develop billing fraud schemes using advances in AI technology.

Data Security

The full scope of the 2024 Change Healthcare data breach included 192.7 million individuals, nearly two-thirds of the U.S. population, making it the largest healthcare data breach in history [18]. The breach exposed names, Social Security numbers, dates of birth, medical information, driver's licenses, passports, and financial payment card data [17]. The Change Healthcare attack is not an isolated incident but part of an accelerating pattern. In 2024, the healthcare sector experienced 742 breach incidents affecting over 276 million records [19]. Since 2020, more than 500 million individuals, exceeding the entire U.S. population, have had healthcare records stolen or compromised [19]. All of these identify data on individual patients can be available on the dark web to further develop claims fraud schemes.

Escalating Threats

These escalating threats—AI-enabled fraud, nation-state exploitation, and systemic

breaches—succeed not because they are undetectable, but because the healthcare transaction infrastructure was never designed to address these challenges.

Several structural deficiencies in the current claims processing environment create the conditions that both enable and conceal fraud at scale:

Fragmented Processing: A single claim often passes through five to seven different, often disconnected, systems before adjudication [20]. Each handoff represents a potential point of data degradation or manipulation, making holistic transaction analysis difficult.

Lack of Standardization: While X12 is the standard format for electronic data interchange (EDI), [21], each payer uses distinct “companion guides” governing specific implementation rules. This variability forces providers to manage dozens of slightly different workflows, creating operational friction that disincentivizes participation by smaller payers and creates loopholes that sophisticated fraudsters exploit [22].

Incomplete Data for Detection: Current fraud detection systems are often trained on public or siloed claims data lacking integration with clinical and practitioner data from Electronic Health Records (EHRs). This incomplete picture limits fraud prediction model accuracy, as payers rarely possess the information needed to algorithmically validate claims [23].

Limitations of Supervised Fraud Models: Supervised fraud detection models depend on court validation of previous fraud cases. If a model incorrectly processes a fraudulent claim as legitimate, similar fraudulent claims may subsequently evade detection.

Legend: ChatGPT completed CMS-1500 professional claim form, the standard form used by non-institutional providers and suppliers to submit claims to Medicare, using synthetic data. Source: National Uniform Claim Committee (NUCC); study authors.

Together, these infrastructure gaps create a compounding vulnerability: fragmented systems prevent the holistic visibility needed to detect fraud, while the absence of standardized, real-time data pipelines makes it impossible to apply the kind of continuous transaction monitoring that has proven effective in financial services. The result is a system structurally locked into post-hoc detection.

The persistence of the reactive model in healthcare creates a systemic barrier to modernization. Payers are hesitant to adopt fully automated claims processing because, without intelligent safeguards, automation could scale fraudulent payments alongside legitimate ones. This creates a critical challenge: payers must either continue with expensive, labor-intensive manual reviews that are overwhelmed by volume, or invest in automated systems equipped to differentiate between legitimate and fraudulent claims in real time.

This paper proposes a strategic and technical framework to address this challenge by adapting proven, multi-layered defense mechanisms from the financial services sector.

Banking Sector Lessons

The financial services industry, shaped by decades of evolving fraud schemes, has developed multi-layered defenses that offer a blueprint to prepare healthcare for the emerging threat of AI-driven fraud mechanisms. These mechanisms have reduced electronic payment fraud losses to a few cents per $100 transacted—a level of integrity to which healthcare can aspire [24].

The banking industry has achieved a 6% to 10% improvement in fraud detection accuracy using AI-powered, real-time systems that analyze transactions in milliseconds [25]. These strategies have also led to the development of novel identity and transactions enterprises that are focused on fraud prevention (Plaid and Stripe).

Insurers in other domains have reported a 30% increase in fraud detection rates and returns on investment (ROI) exceeding 210% within the first year of implementing similar technologies [26]. Deloitte projects that comprehensive AI-driven approaches could save property and casualty insurers between $80 billion and $160 billion by 2032 [27].

Core Banking Anti-Fraud Mechanisms

Robust Customer Identification: Financial institutions use Know Your Customer (KYC) processes to verify identity and credentials of new customers, performing background checks and risk scoring to prevent bad actors from entering the system. This continuous process of identity verification forms the first line of defense against fraud.

Multi-Layer Payment Authorization (“Zero Trust”): No transaction is trusted by default. Layers of verification—including PINs, biometrics, one-time passwords, and geolocation matching—are dynamically applied based on transaction risk profiles.[2] This philosophy assumes every transaction may be fraudulent until proven otherwise through automated verifications.

Continuous Transaction Monitoring (Anti-Money Laundering Analytics): Banks employ

real-time anti-money laundering (AML) analytics to detect suspicious transactions as they occur. Algorithms flag anomalies against expected behavior—such as unusual spending patterns, transactions from high-risk geographies, or rapid fund movements—triggering automated reviews or holds on funds.

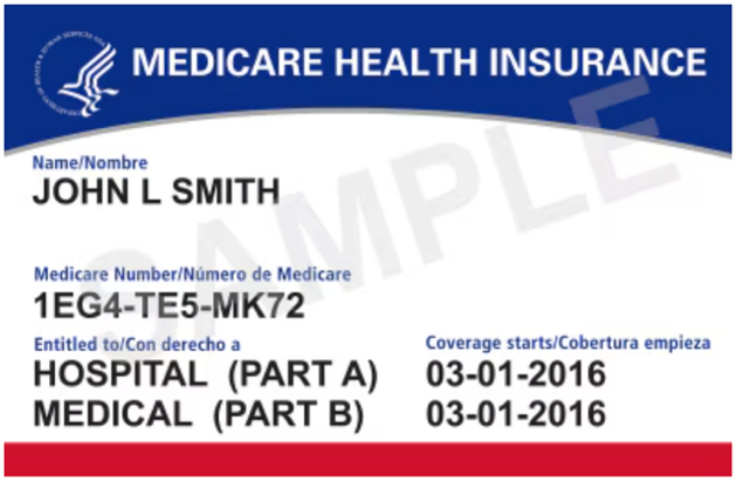
Translation to Healthcare

These principles can be translated to the healthcare ecosystem:

KYC → Member Credentialing

One of the most familiar aspects of the US healthcare system is the need to copy or type the member health plan identification information into a provider registration system. This simple process highlights one of the most critical failures of the current healthcare system: it is based on an analog process of identification where data can easily be copied incorrectly, while data on the patient and the plan (and the plan contract with the provider) are not digitally linked.

The simple contrast between a typical health insurance card and a credit card is illustrative. The health insurance card carries no digital information; it even lacks the most rudimentary digital technology of barcodes.



Legend: Medicare insurance card and Visa card with embedded EMV chip.

Source: https://www.medicare.gov/basics/get-started-with-medicare/using-medicare/your-medicare-card; Bachelot Pierre J-P, CC BY-SA 3.0 https://creativecommons.org/licenses/by-sa/3.0, via Wikimedia Commons

The scale of the gap between banking and healthcare payment security is well documented. The average cost to process a single primary care claim is $20.49 [28] and claims routinely traverse a multi-step adjudication process taking days to months. By contrast, the Visa payment network processes approximately 1,700 transactions per second [29], with AI-powered fraud detection analyzing over 500 features per transaction in real time and preventing $25 billion in fraud in 2019 alone [30]. Banking has developed the EMV (an acronym for Europay, Mastercard, and Visa) chip-enhanced payment security that establishes a unique digital key for each transaction. EMV chip technology reduced counterfeit payment fraud by 76% within three years of adoption [31,32], although this effort pushed fraud to card-not-present transactions [33]. Healthcare's estimated fraud rate of 3% to 10% of total spending, representing $59 billion to $84 billion across the U.S. healthcare system, persists in part because the sector lacks any comparable realtime digital verification mechanism [34].

Member credentialing also faces adoption obstacles that banking did not. Public resistance to digitally coded patient identity is well documented: HIPAA directed HHS to create a unique health identifier for individuals in 1996, but Congress has prohibited funding it since 1999 over privacy objections, and a 2019 House vote to lift the ban stalled in the Senate [35]. National survey data reflect the same sensitivity, with roughly seven in ten individuals reporting concerns about the privacy of their medical records, even as a similar majority supports electronic records and health information exchange [36]. Access is a second obstacle: 91% of U.S. adults own a smartphone and 78% have home broadband, with adoption lowest among older and lower-income adults, the populations most likely to depend on Medicare and Medicaid [37]. Digital member credentials will therefore require clear communication about what data they carry and how they are used, simple non-digital fallback processes for members who cannot or will not adopt them, and visible benefits such as faster eligibility verification at the point of care. Implementation of this solution would have to account for the challenges of vulnerable patient populations, patients with disabilities, and patients that lack full broadband connectivity. Of course, in-person verification of patient identification at the time of a visit is a robust security check.

KYB-Provider Credentialing and Enrollment: Currently, each health plan must develop a provider directory, creating duplication across plans and increasing paperwork burden for providers. A secure KYB transaction environment would establish secure digital credentials (National Provider Identifier [NPI], licensure, billing entity, billing location) for transaction platform entry. Since fraudulent claims submissions is the largest scale risk of AI-generated fraud, KYB efforts likely have the highest return on investment in developing ore robust identity schemes.

Prepayment audits and account verification would establish an additional level of security. Real-time screening of network applicants against the HHS Office of Inspector General's (OIG) List of Excluded Individuals/Entities (LEIE) [38] could further reduce entry of potentially fraudulent entities. The Centers for Medicare and Medicaid Services' (CMS) enhanced screening programs have already deactivated over 660,000 suspect enrollment records from 2011 to 2015, demonstrating the scale of this opportunity [39].

Zero-Trust Payments → Zero-Trust Claims: Payers need systems to validate each transaction as legitimate: verifying the provider, patient, and transaction itself. Banks validate

credit card transactions through EMV (Europay, Mastercard, and Visa) chips using one-time codes. According to Visa, chip-enabled cards reduced card-present counterfeit payment fraud by 76% at chip-enabled merchants within three years of the 2015 U.S. EMV transition [32]. A zero-trust model for healthcare would require that high-risk claims clear additional automated and, when necessary, human checks before payment authorization. The GAO has noted that pre-payment audits are more cost-effective than post-payment recovery efforts [40].

AML Analytics → Claims Pattern Recognition: A digital backbone for the transaction system enables new tools to detect digital fraud, from unexpected surges in claims from a single NPI to ensuring that coded services reflect a provider's practice pattern. Graphical analysis of payer-provider relationships can be used to further enhance the effectiveness of this security layer.

Implementation Framework

This framework adapts the banking sector's three-layer architecture to healthcare fraud prevention, integrating defenses directly into claims workflows

Layer 1: Healthcare Identity Management (KYC/KYB)

The first layer prevents fraudulent actors from entering the system by strengthening provider and patient identity controls.

Enhanced Provider Credentialing: Implement rigorous, automated, continuous credentialing processes including: primary source verification of licenses through API calls to state medical boards; continuous background checks against federal and state exclusion lists (such as the OIG's LEIE, which contains over 660,000 excluded parties); and physical or virtual validation of practice addresses to expose phantom clinics [40]. Providers can be stratified by risk level, with higher-risk entities (e.g., durable medical equipment providers) facing enhanced due diligence.

Digital Provider and Patient ID Authentication: Introduce modern identity-proofing for claim submission, such as multi-factor authentication or digital certificates tied to provider identity. This prevents use of stolen billing credentials. Similarly, strengthening patient identification at the point of service, potentially through biometric-linked insurance cards or secure digital wallets, can reduce identity swapping and medical identity theft. REAL ID

verification is a powerful identity mechanism now in place [41]. Patient identity verification should leverage existing FHIR-based standards to enable interoperable identity tokens that travel with the patient across providers and payers.

Unified Provider Directory Infrastructure: Rather than each health plan maintaining its own provider directory, a centralized or federated provider registry, modeled on the NPPES database but enhanced with real time credentialing status, would reduce duplication, lower administrative burden, and close gaps that enable phantom providers to enter the system undetected. Such infrastructure already has precedent in CMS's Provider Enrollment, Chain, and Ownership System (PECOS), though the current system lacks real time validation capabilities.

Layer 2: Continuous Claim Validation (Zero Trust)

This layer applies the "never trust, always verify" principle to every transaction.

Pre-Payment Review and Smart Adjudication: High-risk claims identified by analytics are automatically routed for pre-payment review or checked against additional data before disbursement. This approach shifts from "pay and chase" to proactive loss prevention.

Machine-Readable Digital Contracts: The technical foundation uses machine-readable standards (X12, FHIR) to encode provider agreements, fee schedules, and benefit plans into computable formats [21]. When a claim is submitted, it can be automatically validated against digital contract terms. For example, a system could automatically deny a second claim for a procedure that a contract specifies is covered only once annually or flag a claim for a service billed by a provider whose digital contract excludes that service. Research indicates that decision tree algorithms embedded in smart contracts can achieve high accuracy in detecting contractual fraud [41].

Clinical Plausibility Engine: Beyond contractual validation, claims should be assessed against clinical logic rules that reflect medical practice standards. For instance, a claim billing for a complex surgical procedure at a facility lacking the appropriate licensure or equipment, or a claim for an office visit submitted on the same day as an inpatient admission, represents clinically implausible activity that can be flagged automatically. Integration with EHR data, where available and permissible under HIPAA's "minimum necessary" standard, would further

strengthen clinical validation without requiring manual chart review. Risk stratification engines would assign dynamic risk scores to each claim based on provider history, claim characteristics, patient demographics, geographic patterns, and clinical plausibility. Only claims exceeding a configurable risk threshold would require enhanced scrutiny, allowing the vast majority of legitimate claims to process at current or faster speeds.

Layer 3: Advanced Analytics and Risk Scoring (AI/ML)

This layer deploys advanced analytics to detect anomalous behavior in real time.

Predictive Modeling and AI: ML models trained on historical claims data generate real-time fraud risk scores for each new claim. High-risk claims are automatically flagged for intervention. These models continuously learn from confirmed fraud cases, adapting to new schemes and improving accuracy while reducing false positive rates over time. CMS's Fraud Prevention System has demonstrated greater than 3-to-1 ROI [42].

Graph and Outlier Analysis: Graph analytics visualizes and analyzes complex, often hidden networks connecting providers, patients, claims, and locations. This technique effectively uncovers collusion, such as kickback rings or organized fraud using multiple shell companies linked to a common address [43].

Collaborative Defense Network: The framework leverages a federated data model inspired by the Healthcare Fraud Prevention Partnership (HFPP). As of 2016, HFPP represented over 65% of covered lives in the U.S. and has since grown to over 300 partner organizations as of 2023 [44][45]. Participants contribute anonymized fraud pattern data and gain access to collective intelligence, enabling smaller payers to access enterprise-grade detection capabilities and larger payers to identify schemes spanning multiple networks.

The framework proposes expanding this federated model to include real time threat intelligence sharing, analogous to the financial sector's FS-ISAC (Financial Services Information Sharing and Analysis Center), enabling participating organizations to disseminate indicators of compromise as they emerge rather than in periodic retrospective reports. Network analysis can detect patterns that evade individual claim review, such as patient sharing rings, unusually concentrated referral patterns, and geographic clustering of high-cost services at addresses

associated with minimal physical infrastructure.

Measuring Success

Effective implementation requires clearly defined performance metrics tracked across all three layers. Layer 1 (KYC) metrics include time to credentialing for new providers, percentage of enrollment applications flagged and confirmed as fraudulent, and reduction in phantom provider entries compared to baseline. Layer 2 (Zero Trust) metrics include percentage of claims validated against digital contract terms in real time, reduction in contractual fraud such as duplicate billing and out of scope services, and claim processing speed relative to pre implementation benchmarks. Layer 3 (AI/ML) metrics include fraud detection rate (true positive rate), false positive rate and its trend over time, dollar value of fraud prevented per dollar invested, and average time from claim submission to fraud identification. System wide metrics include overall improper payment rate for participating organizations versus nonparticipants, provider satisfaction and administrative burden metrics, and return on investment at each phase of implementation. These metrics should be independently audited and published to build the evidence base for broader adoption and to inform regulatory policy.

Risk Considerations

An aggressive fraud prevention program must be implemented thoughtfully to mitigate potential risks. Each layer of the framework carries potential unintended consequences for beneficiaries, providers, and payers. Acknowledging these risks explicitly and building mitigation into the design is as important as the fraud prevention itself. Provider risks are addressed in Section 6.3, payer risks in Section 6.4, and beneficiary risks in Section 6.5.

Regulatory and Legal Risks

All activities must comply with privacy laws including HIPAA and various state regulations, as well as prompt pay laws. While HIPAA permits use of protected health information (PHI) for fraud detection, a “minimum necessary” standard must be enforced. Pre-payment reviews must avoid violating claim payment timelines. To prevent False Claims Act liability, any automated denial process must include human oversight for ambiguous cases, a principle now required by law in states including California [46].

Cybersecurity and System Integrity

New analytics platforms and data-sharing networks expand cybersecurity risk. All systems require robust end-to-end encryption, multi-factor access controls, and regular penetration testing. If utilized, smart contracts require intensive code security audits to prevent exploitation of vulnerabilities [21].

False Positives and Provider Relations

The most significant operational risk is flagging legitimate claims (false positives), which can delay payments and damage provider relationships. This risk requires active management through:

- **Calibrating AI Thresholds:** Start with conservative settings flagging only high-confidence outliers, tuning to minimize provider friction initially while allowing threshold adjustment as models improve.
- **Transparent and Rapid Processes:** Establish clear, expedited appeals processes for providers to contest flagged claims.
- **Provider Education and Collaboration:** Communicate program goals and mechanics to network providers, framing it as collaborative effort to protect system integrity and accelerate payment for legitimate claims.

Pre-payment review also carries cash flow and workload consequences for providers. Physician practices already devote substantial resources to payer review processes: in the 2025 American Medical Association survey, physicians completed an average of 40 prior authorization requests per week, consuming roughly 13 hours of physician and staff time, and more than a quarter reported that these review processes led to serious adverse events for patients [47]. Fraud screening must not impose a comparable burden on honest providers. Risk-based exemptions offer one proven mitigation: under the Texas “gold card” law, providers whose requests are consistently approved earn exemptions from preauthorization review, a model that could similarly exempt consistently low-risk providers from enhanced claim scrutiny [48]. Clean claims from exempt providers should be paid within contractual and statutory prompt pay timelines, with flagging criteria disclosed to network providers and screening performance published as described in Section 5.4.

Implementation, Operational, and Payer Risks

AI system success depends on data quality. The project requires strong interdepartmental governance ensuring coordination between IT, compliance, claims, and provider relations. A phased rollout, starting with less complex claim types, and continuous performance monitoring are essential. For payers, the principal risks are financial and reputational: large technology investments may underdeliver if data quality is poor, and poorly calibrated models can create regulatory exposure and erode provider and member trust. Phased rollouts with published performance metrics (Section 5.4) help contain these risks. Payers also face model risk that is specific to an adversarial setting: fraud perpetrators actively probe and adapt to detection systems, so model performance degrades without continuous retraining, monitoring, and validation [49]. Governance should therefore include periodic independent model audits, documented retraining cycles, and fallback to human review when model confidence is low.

Risks to Beneficiaries

Beneficiaries ultimately bear the consequences of fraud prevention errors. A legitimate claim wrongly flagged can mean a delayed authorization, an unexpected bill, or an interruption in care. Any adverse action that affects access to care should therefore require human review before it takes effect, and patients should be held harmless while payer and provider disputes are resolved. Beneficiaries affected by a flagged claim should receive plain language notice and a simple appeal channel; payer review processes are already associated with care delays and treatment abandonment [50]. Stronger identity verification can also create access barriers for beneficiaries who lack smartphones, reliable internet access, or digital literacy; simple non-digital pathways to verify identity and receive care must be preserved. Aggregating identity, clinical, and financial data for fraud analytics raises privacy concerns beyond regulatory compliance, so programs should collect the minimum data necessary, disclose plainly how the data are used, and submit to independent audits. Finally, detection models should be monitored for disparate impact to ensure fraud controls do not systematically burden specific patient populations, such as those in underserved areas where documentation practices may differ.

Conclusion

The healthcare industry is increasingly vulnerable to threats of AI-enabled digital fraud, especially in the claims financial transaction system. Addressing this challenge will require

transformation of the payment system to preserve the integrity of payment processes against this emerging risk by adopting a proactive prevention framework inspired by the financial services sector. A three-layer defense system built on rigorous identity management, zero-trust validation, and AI-powered analytics, can transform claims processing from a reactive, cost-intensive function into a proactive platform. Further leveraging machine-readable contracts as the foundation, payers can automate fraud detection at the point of submission, creating potential ROI that aligns financial incentives with technological modernization. Realizing these benefits will require managing the framework's risks to beneficiaries, providers, and payers with the same rigor as the effort applied to fraud prevention.

References

[1] U.S. Government Accountability Office. (2024, November). Medicare and Medicaid: Additional Actions Needed to Enhance Program Integrity and Save Billions (GAO-24-107487). https://www.gao.gov/products/gao-24-107487

[2] U.S. Department of Health and Human Services & Centers for Medicare & Medicaid Services. (2025). Trump Administration Prioritizes Affordability, Announcing Major Crackdown on Health Care Fraud [Press release]. https://www.cms.gov/newsroom/press-releases/trump-administration-prioritizes-affordability-announcing-major-crackdown-health-care-fraud

[3] Centers for Medicare & Medicaid Services. (2016). Common Types of Health Care Fraud [Fact sheet]. https://www.cms.gov/files/document/overviewfwacommonfraudtypesfactsheet072616pdf

[4] U.S. Department of Justice. (2025, June 30). National Health Care Fraud Takedown Results in 324 Defendants Charged in Connection with Over $14.6 Billion in Alleged Fraud. https://www.justice.gov/opa/pr/national-health-care-fraud-takedown-results-324-defendants-charged-connection-over-146

[5] National Health Care Anti-Fraud Association. (n.d.). The Challenge of Health Care Fraud. https://www.nhcaa.org/tools-insights/about-health-care-fraud/the-challenge-of-health-care-fraud/

[6] California Health Advocates. (n.d.). Medicare Checks for Fraud Before Paying. https://cahealthadvocates.org/medicare-checks-for-fraud-before-paying/

[7] Sumsub. (2025). Identity Fraud Report 2024–2025. https://sumsub.com/blog/guides-reports/identity-fraud-report-2024-2025/

[8]). Insurance fraud increased by 19% from synthetic voice attacks in 2024. Fierce Healthcare. https://www.fiercehealthcare.com/payers/insurance-fraud-increased-19-synthetic-voice-attacks-2024

[9] U.S. Department of Justice. (2025, August 20). Troy Health, Inc. Enters Non-Prosecution Agreement and Admits to Fraudulently Enrolling Medicare Beneficiaries and Identity Theft. https://www.justice.gov/opa/pr/troy-health-inc-enters-non-prosecution-agreement

[10] Centers for Medicare & Medicaid Services. (2025, June). CMS Prevented Over $4 Billion in Fraudulent Claims. Referenced in the DOJ 2025 National Health Care Fraud Takedown announcement.

[11] CISA, FBI, and DC3. (2024, August 28). Iran-based Cyber Actors Enabling Ransomware Attacks on US Organizations (Advisory AA24-241A). https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-241a

[12] CISA, FBI, and U.S. Department of the Treasury. (2022, July 6). North Korean State-Sponsored Cyber Actors Use Maui Ransomware to Target the Healthcare and Public Health Sector (Advisory AA22-187A). https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-187a

[13] FBI and CISA. (2020, July 22). People’s Republic of China (PRC) Targeting of COVID-19 Research Organizations. https://www.cisa.gov/news-events/alerts/2020/07/22/peoples-republic-china-prc-targeting-covid-19-research-organizations

[14] National Cyber Security Centre (UK). (2020, July 16). Advisory: APT29 Targets COVID-19 Vaccine Development.

https://www.ncsc.gov.uk/news/advisory-apt29-targets-covid-19-vaccine-development

[15] U.S. Department of Health and Human Services, Health Sector Cybersecurity Coordination Center (HC3). (2024). Healthcare Sector Ransomware Trend Analysis [HC3 Threat Brief].

[16] Health Information Sharing and Analysis Center (Health-ISAC). (2025, February). 2025 Health Sector Cyber Threat Landscape Report. https://health-isac.org/health-isac-2025-health-sector-cyber-threat-landscape/

[17] Kang C. Google Says Chinese Cybercrime Group Used Its A.I. in Scams. NY Times. June 12, 2026. https://www.nytimes.com/2026/06/12/technology/google-lawsuit-china-ai-scams.html

[18] U.S. Department of Health and Human Services, Office for Civil Rights. (2025, July 31). Change Healthcare Updated Breach Notification. https://www.hhs.gov/hipaa/for-professionals/special-topics/change-healthcare-cybersecurity-incident-frequently-asked-questions/index.html

[19] HIPAA Journal. (2025). Healthcare Data Breach Statistics. https://www.hipaajournal.com/healthcare-data-breach-statistics/

[20] Renfrow, J. (2019, April 10). 90% of claim denials are avoidable with help of tech tools. Fierce Healthcare. https://www.fiercehealthcare.com/payer/waystar-90-claim-denials-are-avoidable-help-technology

[21] Lauren Bilbo, Elman Amador Medina, Dashiell Miner, David Scheinker, Stefanos Zenios, Kevin Schulman. 2026. Developing a Computable Payer-Provider Contract. *Health Management, Policy and Innovation* (www.HMPI.org), Volume 11, Issue 1.

[22] Invene. (2025, August). 270/271, 835, 837: Decoding the 9 Key Healthcare EDI Transactions. https://www.invene.com/blog/demystifying-healthcare-edi-the-9-critical-transactions-explained

[23] Straits Research. (2024). Healthcare Fraud Analytics Market Size, Global Trends, Share, Forecast to 2033. https://straitsresearch.com/report/healthcare-fraud-analytics-market

[24] The Nilson Report. (2022, December). Card Fraud Losses Reach $32.40 Billion (Issue 1230). Referenced in: Payments Dive. (2023, January 5). Card industry's fraud-fighting efforts pay off. https://www.paymentsdive.com/news/card-industry-fraud-fighting-efforts-pay-off-nilson-report-credit-debit/639675/

[25] Flinders, M., Smalley, I., & Schneider, J. (n.d.). AI fraud detection in banking. IBM Think. https://www.ibm.com/think/topics/ai-fraud-detection-in-banking

[26] Aloa. (2025, August 28). How AI is changing fraud detection in insurance. https://aloa.co/blog/ai-is-changing-fraud-insurance

[27] Deloitte. (2025, April 24). Using AI to fight insurance fraud. Deloitte Insights. https://www.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2025/ai-to-fight-insurance-fraud.html

[28] Tseng, P., Kaplan, R. S., Richman, B. D., Shah, M. A., & Schulman, K. A. (2018). Administrative Costs Associated with Physician Billing and Insurance-Related Activities at an Academic Health Care System. JAMA, 319(7), 691–697.

[29] Gillai, B., & Mendelson, H. (2020, November). Creating Value with Blockchain: A Value Chain Management Perspective. Stanford Graduate School of Business. https://www.gsb.stanford.edu/faculty-research/publications/creating-value-blockchain-value-chain-management-perspective

[30] Dar, N., et al. (2021, April 14). Technology-Led Shifts and Opportunities in Card-Based Payments. McKinsey & Company. https://www.mckinsey.com/industries/financial-services/our-insights/banking-matters/technology-led-shifts-and-opportunities-in-card-based-payments

[31] Bowman, C. (2022, February 9). What Are EMV Chips and Do They Make Credit Cards More Secure? CNET.

https://www.cnet.com/personal-finance/credit-cards/what-are-credit-card-chips-and-are-they-more-secure/

[32] Visa. (2018, May). Counterfeit fraud at chip-enabled merchants down 76 percent. Reported in: American Bankers Association. (2018, May). Visa: Counterfeit Fraud Drops 76 Percent at Chip-Enabled Merchants. ABA Banking Journal. https://bankingjournal.aba.com/2018/05/visa-counterfeit-fraud-drops-76-percent-at-chip-enabled-merchants/

[33] Hayashi F. New Data on Card-Present and Card-Not-Present Fraud Rates in the United States. Federal Reserve Bank of Kansas City. February 25, 2026. https://www.kansascityfed.org/research/payments-system-research-briefings/new-data-on-card-present-and-card-not-present-fraud-rates-in-the-united-states/

[34] Shrank, W. H., et al. (2019). Waste in the US Health Care System: Estimated Costs and Potential for Savings. JAMA, 322(15), 1501–1509.

[35] Pifer, R. (2019, June 14). House Votes to Overturn Decades-Old Ban on National Patient Identifier. Healthcare Dive. https://www.healthcaredive.com/news/house-votes-to-overturn-decades-old-ban-on-national-patient-identifier/556859/

[36] Patel, V., Hughes, P., Savage, L., & Barker, W. (2015, June). Individuals' Perceptions of the Privacy and Security of Medical Records (ONC Data Brief No. 27). Office of the National Coordinator for Health Information Technology. https://www.healthit.gov/sites/default/files/briefs/oncdatabrief27june2015privacyandsecurity.pdf

[37] McClain, C. (2026, January 8). What We Know About Internet Use, Smartphone Ownership and Digital Divides in the U.S. Pew Research Center. https://www.pewresearch.org/short-reads/2026/01/08/internet-use-smartphone-ownership-digital-divides-in-u-s/

[38] U.S. Department of Health and Human Services, Office of Inspector General. List of Excluded Individuals/Entities (LEIE) [updated monthly]. https://oig.hhs.gov/exclusions/

[39] U.S. Government Accountability Office. (2016, November). Medicare: Initial Results of Revised Process to Screen Providers and Suppliers, and Need for Objectives and Performance Measures (GAO-17-42). https://www.gao.gov/products/gao-17-42

[40] U.S. Government Accountability Office. (2016, April). Medicare: Claim Review Programs Could Be Improved with Additional Prepayment Reviews and Better Data (GAO-16-394). https://www.gao.gov/products/gao-16-394

[41] Amponsah, A. A., Adekoya, A. F., & Weyori, B. A. (2022). A novel fraud detection and prevention method for healthcare claim processing using machine learning and blockchain technology. Healthcare Analytics, 2, Article 100034. https://www.sciencedirect.com/science/article/pii/S2772662222000534

[42] U.S. Department of Health and Human Services, Office of Inspector General. (2016, September). The Fraud Prevention System's Adjustment Factors (Report No. A-01-15-00510). https://oig.hhs.gov/oas/reports/region1/11500510.pdf

[43] Chan, T. (2024, January 16). Data visualization and AI for healthcare fraud detection. Cambridge Intelligence Blog. https://cambridge-intelligence.com/medical-fraud-detection/

[44] Centers for Medicare & Medicaid Services. (2016). The Health Care Fraud and Abuse Control Program Protects Consumers and Taxpayers by Combating Health Care Fraud [Fact sheet]. https://www.cms.gov/newsroom/fact-sheets/health-care-fraud-and-abuse-control-program-protects-consumers-and-taxpayers-combating-health-care-0

[45] U.S. Department of Health and Human Services, Office of Inspector General and Department of Justice. (2024, December 6). Health Care Fraud and Abuse Control Program Annual Report for Fiscal Year 2023. https://oig.hhs.gov/reports/all/2024/health-care-fraud-and-abuse-control-program-report-fiscal-year-2023/

[46] California Senate Bill 1120 (Becker). (2024). Health Care Coverage: Utilization Review. Signed September 28, 2024; effective January 1, 2025. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB1120

[47] American Medical Association. (2026). 2025 AMA Prior Authorization Physician Survey. https://www.ama-assn.org/system/files/prior-authorization-survey.pdf

[48] Texas Department of Insurance. (2025). Preauthorization Exemptions (House Bill 3459, 87th Legislature, 2021; updated by House Bill 3812, 89th Legislature, 2025). https://www.tdi.texas.gov/health/hb3459.html

[49] Lunghi, D., Simitsis, A., Caelen, O., & Bontempi, G. (2023). Adversarial Learning in Real-World Fraud Detection: Challenges and Perspectives. Second ACM Data Economy Workshop (DEC '23). https://doi.org/10.1145/3600046.3600051

[50] Pickern JS. Prior authorizations and the adverse impact on continuity of care. Am J Manag Care. 2025 Apr;31(4):163-165.

[1] Deepfakes are synthetic audio, video, or images created with artificial intelligence to convincingly impersonate real people.

[2] Some of these security approaches may be challenging for elderly patients or patients with dementia.