# **Usage of Transformers in U.S. Healthcare Payment Integrity**

## Introduction

The U.S. healthcare system processes over 5 billion claims annually, with improper payments estimated to exceed **$100 billion per year**. As the industry shifts toward **value-based care (VBC)**, ensuring payment integrity has become both more complex and more critical. Traditional statistical models and rule-based systems are increasingly inadequate in detecting nuanced patterns of fraud, waste, and abuse (FWA). This research proposes the use of **transformer-based deep learning models** to enhance payment integrity through advanced natural language processing (NLP) and anomaly detection.

## Research Problem

Despite the availability of large-scale healthcare data, current quantitative methods face limitations in:

* Processing unstructured data (e.g., clinical notes, EHRs)
* Understanding contextual nuances in medical necessity
* Detecting real-time fraud with low false positives

Recent studies show:

* **75% of leading healthcare companies** are experimenting with or planning to scale generative AI across the enterprise
* **46% of U.S. healthcare organizations** are in the early stages of implementing generative AI
* **40% of U.S. physicians** are ready to use generative AI at the point of care in 2025
* The **payment integrity industry** is valued at **$9 billion**, with AI poised to transform it

## Objectives

* To develop and evaluate transformer-based models for detecting anomalies in healthcare claims.
* To quantify improvements in fraud detection accuracy, precision, and recall over traditional models.
* To assess the scalability and interpretability of these models in real-world payer environments.

## Methodology

* **Data**: De-identified claims data, EHRs, and provider notes.
* **Modeling**: Fine-tuning transformer models (e.g., ClinicalBERT, RoBERTa) for:
  + Named Entity Recognition (NER)
  + Semantic similarity for duplicate claim detection
  + Classification of high-risk claims
* **Quantitative Techniques**:
  + ROC-AUC, confusion matrix analysis
  + Time-series anomaly detection
  + Statistical validation using bootstrapping and cross-validation

## Significance

This research bridges **quantitative analysis** and **deep learning**, offering a novel approach to a high-impact problem. It aligns with national priorities on healthcare cost containment and fraud prevention and contributes to the growing field of **explainable AI in healthcare**.

## Recent Developments

* The **CMS ACO Primary Care Flex Model**, launched in 2025, emphasizes AI-driven care coordination and payment optimization
* **29% of healthcare leaders** report already using generative AI tools, while **43% are testing them**
* AI is increasingly used to extract value from unstructured EHR data, a key enabler for VBC and payment integrity

## Expected Outcomes

* A validated transformer-based framework for payment integrity.
* Quantitative benchmarks comparing traditional vs. transformer models.
* Policy and implementation recommendations for healthcare payers.