Kelvin Open Science Publishers
Connect with Research Community

Research Article Volume 1 | Issue 1

KOS Journal of AIML, Data Science, and Robotics
https://kelvinpublishers.com/journals/aiml-data-science-robotics.php

AI-Driven Risk Stratification Models for Medicaid: Algorithms, Bias, and Validation Challenges

Anand Laxman Mhatre

Senior Program Manager Accenture, State of Texas Austin, TX, USA

***Corresponding author:** Anand Laxman Mhatre, Senior Program Manager Accenture, State of Texas Austin, TX, USA

Received: May 12, 2025; **Accepted:** May 24, 2025; **Published:** May 26, 2025

Citation: Anand LM. (2025) AI-Driven Risk Stratification Models for Medicaid: Algorithms, Bias, and Validation Challenges. *KOS J AIML, Data Sci, Robot.* 1(1): 1-3.

1. Abstract
AI risk stratification models can play a significant role in creating patient risk scores, allowing Medicaid managers to project care delivery costs and allocate resources more efficiently. However, these models can be biased and opaque, hampering their effectiveness in risk stratification. This document explores algorithms that can be deployed in Medicaid risk stratification and their vulnerabilities.

2. Keywords
Risk stratification, Artificial intelligence, Medicaid, Bias, Algorithms

3. Introduction
As the era for value-based care ushers in, AI risk stratification models will allow Medicaid programs to identify high-risk patients, allocate resources efficiently, and improve health outcomes. These models can quickly and accurately generate members' risk scores, allowing Medicaid management to accurately project costs and optimally allocate resources. Although these models are essential in delivering Medicaid services, they have a fair share of issues. They tend to be biased and can be opaque as they become more sophisticated [1].

4. Algorithms
According to Syed, AI risk stratification models work by leveraging member-specific variables. They use metrics such as social determinants of health and prevailing health indicators to calculate probabilities of complications, disease progression, and treatment outcomes [2]. These models rely on machine learning algorithms that have distinct capabilities and trade-offs. The top algorithms used in Medicaid AI risk stratification models include;

4.1. Logistic regression
Logistic regression is one of the most used algorithms in risk stratification. It is ideal for measuring the probability of binary outcomes. For example, it determines whether patients are likely to be readmitted or not. The algorithm is straightforward and is valued for its simplicity and ease of interpretability. However, it's only effective for linear data.

4.2. Gradient boosting machines
Constitutes algorithms such as XGBoost, LightGBM, and CatBoost. These algorithms create predictive models by improving previous iterations. They have better predictive accuracy and are typically used in complex risk stratification applications like forecasting chronic disease progression, frequency of emergency room visits, and cost estimates. These algorithms are complex to interpret,

necessitating tools such as SHAP values.

4.3. Clustering algorithms

They include algorithms such as k-means and hierarchical clustering. They are used to segment populations based on shared features that are not readily identifiable. For example, they can group populations based on similar disease profiles or healthcare utilization patterns using attributes hidden in data.

4.4. Neural networks

Neural networks such as deep learning models are excellent in modeling high-dimensional data. This capability makes them perfect for dealing with large-scale time series, unstructured health records. Neural networks can autonomously create member risk scores and classify them accordingly. However, this ability to function autonomously renders them susceptible to the 'black box' phenomenon.

4.5. Survival analysis models

Include algorithms such as Cox Proportional Hazards and DeepSurv. These algorithms measure time-to-event outcomes, such as time until hospitalization or death. They use censored data to provide time-based risk estimates. They are ideal for determining the necessity of proactive interventions.

4.6. Natural language processing

In risk stratification, these models are primarily used to complement other models. Their primary role is to extract and analyze unstructured data from documents such as medical notes, patient communications, and social work assessments. Such tools include ClinicalBERT, which can extract health indicators from documents and enrich risk models with social and behavioral determinants.

4.7. Decision trees and random forests

These intuitive models split data into branches based on their features. Models based on decision trees split Medicaid members into branches based on social determinants of health and prevailing health conditions. Random trees are easy to implement but can be prone to overfitting problems. This problem can be mitigated by employing random forests, which work by averaging the predictions of many trees built on different subsets of data.

Algorithm	**Use Case in Healthcare**	**Limitations**
Logistic Regression	Predicting binary outcomes (readmission)	Struggles with complex data
Decision Trees	Classifying patient risk levels	Susceptible to overfitting
Random Forests	Predicting high-risk patients,	Slower with large datasets
Gradient Boosting	Cost prediction, ED visits, disease onset	Less transparent, needs explanation tools
Neural Networks	Modeling complex EHR or time-series data	Black-box problem
Clustering	Segmenting population into risk groups	Requires tuning, sensitive to initial settings
NLP Models	Extracting risk factors from clinical notes	Complex models, requires pre-processing
Survival Analysis	Predicting time to hospitalization, death	Traditional models assume proportional hazards

5. Bias in Risk Stratification Models

One of the leading issues of using AI in risk stratification is the risk of bias. When models are not trained properly and monitored, they can generate skewed outcomes that favor some groups at the expense of others. AI bias can be caused by factors such as

- ***Biased data***: AI algorithms are trained by training data sets. When the training data consists of biases, partialities will inadvertently show up when the algorithm is deployed in real working environments [3].
- ***Label bias***: Using unfavourable attributes to classify risks can cause bias. For example, using cost or utilization as proxies for health needs can underestimate the healthcare needs of disadvantaged groups.
- ***Algorithmic bias***: This happens when the design and parameters of the model inadvertently introduce biases. For example, the algorithm may be designed to prioritize certain features not present in disadvantaged groups, which may lead to discriminatory outcomes.
- ***Human decisions bias***: Human bias from developers and people operating the systems may slip into the models. For example, if an AI model learns that their outputs are often modified, moving members of certain populations to specific stratification, it may adjust its subsequent outputs accordingly.

Biases can exacerbate existing societal inequities, denying certain groups equitable access to Medicaid services. Biases can also reinforce societal stereotypes, incorrectly labeling certain groups as more susceptible to specific health conditions. Biases in Medicaid risk scoring can result in suboptimal health plans for disadvantaged groups, exasperating health disparities. Biases in AI risk stratification models can be mitigated using the following strategies.

- ***Data pre-processing***: This entails cleaning and balancing the training data to remove imbalance before using it to train AI models [4].
- ***Auditing and transparency***: This involves assessing AI model outputs and adjusting them to remove bias. Developers can ensure transparency by explaining how the models make decisions.
- ***Fairness Awareness algorithms***: This approach involves coding rules and guidelines into the algorithms to avoid algorithmic biases.
- ***Diversified development teams***: Diversified development teams are more likely to build risk stratification models that are less biased.

5.1. Validation challenges

AI algorithms are complex and dynamic, making validation a challenge. Some of the areas where validation challenges are persistent include:

- ***Data volume***: Training Medicaid risk stratification models requires large quantities of data. Ensuring the quality and representativeness of this data can be a challenge.
- ***Data drift***: AI algorithms are always learning. When the data changes over time, the accuracy of decisions may also change.
- ***Black-box problem***: Some algorithms may autonomously optimize their decision-making parameters to the extent that users and developers no longer understand how they make decisions. In such a case, it may be problematic to validate the accuracy of the models.
- ***Generalization***: Medicaid risk stratification models developed for a particular population may not perform optimally on another population due to demographic differences.
- ***Measuring fairness***: There are no standard metrics for measuring fairness.

While addressing validation challenges can be difficult, leveraging various strategies can reduce these issues' impact. The training data should come from multiple sources, the system should be audited routinely for bias, and community stakeholders should be involved in evaluating risk stratification models before launch. Also, explainable AI should be used in development, and the models should be optimized continuously to ensure fairness and accuracy are maintained.

6. Conclusion

AI holds immense potential in Medicaid risk stratification. However, failure to address issues such as bias and validation challenges can result in these tools reinforcing disparities that they aim to resolve. This piece of writing recommends a multidisciplinary approach – community input, diverse technical expertise, and oversight – as the best way to build Medicaid AI models that are equitable and effective.

7. References

1. Fainman AA. (2020) opaque AI. *Fourth Industrial Revolution*. 44.
2. Mohsin SN, Gapizov A, Ekhator C, et al. (2023) The role of artificial intelligence in prediction, risk stratification, and personalized treatment planning for congenital heart diseases. *Cureus*. *15*(8).
3. SAP. (2024) What is AI bias?
4. Bhandari A. (2023) Bias in AI: A comprehensive examination of factors and improvement strategies. *Int. J. Comput. Sci. Eng. 10*(6): 9-14.