Ethics in Personalized Medicine: Analysis of AI with TCGA Dataset

Fakii Mohammed, Andrew Ogembo, Chiboniso Nyoni, Peterson Kagiri
Group 67 AI Software Engineers

July 2025

Abstract

This 300-word analysis by Group 67 AI Software Engineers examines ethical challenges in using AI with the Cancer Genomic Atlas (TCGA) dataset for personalized medicine. It identifies biases in treatment recommendations, particularly due to underrepresentation of ethnic groups, and proposes fairness strategies to ensure equitable outcomes.

1 Ethical Analysis

AI-driven treatment recommendations using the TCGA dataset risk biases that undermine equitable healthcare. A primary concern is the underrepresentation of ethnic groups in the dataset. TCGA, while comprehensive, predominantly includes genomic data from Caucasian patients, with limited representation of African, Asian, or Indigenous populations. This skew can lead AI models to generate inaccurate or ineffective treatment recommendations for underrepresented groups. For example, genetic variations in cancer susceptibility, such as BRCA mutations, differ across ethnicities, and a biased model may overlook these, leading to suboptimal treatments or misdiagnoses, exacerbating health disparities.

Socioeconomic factors further compound biases, as TCGA data often comes from patients with access to advanced healthcare, excluding lower-income or rural populations. This can result in models favoring treatments inaccessible to marginalized groups. Additionally, algorithmic biases may arise from training on imbalanced datasets, where overfitting to majority groups reduces generalizability.

To address these, Group 67 proposes fairness strategies. First, diversify training data by integrating datasets from global health initiatives, like the African Genome Variation Project, to include varied ethnic profiles. Second, implement bias audits using metrics like demographic parity to detect disparities in model predictions. Third, adopt explainable AI techniques, such as SHAP, to make treatment rationales transparent, enabling clinicians to verify fairness. Finally, engage communities through participatory design to align AI models with diverse patient needs. These strategies, championed by Fakii Mohammed, An-

drew Ogembo, Chiboniso Nyoni, and Peterson Kagiri, ensure AI in personalized medicine promotes equity and trust (300 words).

References

[1] TCGA Dataset, `https://www.kaggle.com/datasets/tcga`.

[2] IEEE Ethically Aligned Design, `https://standards.ieee.org`.