**Task 3: Ethics in Personalized Medicine**  
**Dataset: Cancer Genomic Atlas (TCGA)**

Artificial Intelligence (AI) is increasingly used in personalized medicine to recommend cancer treatments based on genomic data. While this promises more targeted and effective care, it also raises ethical concerns—particularly regarding **bias and fairness**.

**1. Potential Biases**

A major issue with using AI on datasets like TCGA is **population underrepresentation**. Most genomic data in TCGA come from individuals of European ancestry. This skews the training data, making AI models less accurate when applied to **ethnic minorities**, such as African, Asian, or Indigenous populations. Consequently, these groups may receive less effective treatment recommendations, potentially worsening health disparities.

Another concern is **data incompleteness**. Many datasets focus on genetic information while overlooking **social determinants of health**—such as income, environment, and healthcare access—which also influence treatment outcomes. This limits the context in which AI makes recommendations, leading to oversimplified or biased results.

**2. Fairness Strategies**

To reduce bias, a key step is to **diversify training data**. This involves expanding TCGA and other datasets to include **genomic and clinical data from a wider range of populations**, especially those historically underrepresented. International collaboration and inclusive recruitment strategies can help achieve this.

Secondly, implement **algorithm audits** to evaluate how model performance varies across demographic groups. If disparities are found, adjustments should be made using techniques like **re-weighting**, **fairness constraints**, or **bias mitigation algorithms**.

Lastly, AI systems should be made **transparent and interpretable** so healthcare providers can understand, question, and refine treatment suggestions. This supports accountability and helps build trust, especially among marginalized communities.