models across populations. From a statistical genetics perspective, cross-population inequity can be understood as inconsistency in the statistical target (estimand) represented by the predictive function across different data domains, leading to systematic performance bias and decision risk.

4.1 Inequity in cross-population prediction

A large body of evidence indicates that PRS/PGS trained predominantly on European populations exhibit substantially reduced predictive performance in non-European populations (approximately 40-80% decline). This reduction is primarily driven by differences in allele frequency spectra and linkage disequilibrium (LD) structure, sampling bias in GWAS discovery and effect estimation, and inconsistencies in phenotype definition and measurement. These factors collectively result in decreased explanatory power, calibration bias, and unstable decision thresholds during extrapolation, thereby increasing misclassification risk and leading to inequitable allocation of resources (Duncan et al., 2019; Martin et al., 2019; Zhang et al., 2023).

From a statistical inference perspective, this phenomenon can be understood as arising from systematic differences across populations in linkage disequilibrium structure, allele frequency spectra, and effect size distributions, which lead to a misalignment between the estimand defined in the training data and the target of prediction in the external population-commonly referred to as estimand mismatch.

Among different populations, individuals of African ancestry and admixed populations are particularly affected, largely due to their higher genetic diversity and the relative lack of representative LD reference panels and functional annotation resources. As a result, tag SNPs are less able to reliably proxy underlying causal variants (Ding et al., 2023; Kachuri et al., 2024).

Similar issues are observed in crop breeding systems. Differences in subspecies structure, ecological population stratification, and the target population of environments (TPE) can lead to failure of PGS models trained on elite germplasm or specific environments when applied to local varieties or marginal ecological conditions. Gene-environment interactions (G×E) further amplify these discrepancies, reducing selection efficiency under environmental extrapolation and potentially leading to systematic neglect of important genetic signals relevant to smallholder farming systems or marginal environments, thereby creating forms of “hidden inequity” (Sima et al., 2024).

4.2 Ethical and societal considerations

Overinterpretation of PRS as a deterministic measure of an individual’s “genetic destiny” may lead to distorted risk perception, stigmatization, and reinforcement of social stereotypes. When predictive performance and decision thresholds are not comparable across populations, the direct application of PRS in screening or intervention may result in unequal distribution of healthcare resources, thereby exacerbating existing health disparities (Martin et al., 2019; Lewis and Green, 2021; Andreoli et al., 2024). In addition, issues related to data governance and privacy are critical, including risks of re-identification, policies for returning results to participants, and mechanisms for dynamic informed consent. These considerations require the integration of ethical evaluation throughout the entire research and implementation lifecycle (Adeyemo et al., 2021). In sensitive domains such as psychiatric traits, particular caution is needed in interpretation and communication to avoid reinforcing genetic determinism or misleading the public (Murray et al., 2020; Chapman, 2022).

In agricultural and breeding contexts, ethical challenges are more closely associated with structural imbalances in technology deployment. For example, recommendations based solely on large-scale data from a single environment may disadvantage niche ecological systems or smallholder farmers, and may, over time, reduce genetic diversity and compromise the resilience of food systems (Sima et al., 2024). Furthermore, access to and utilization of international germplasm resources and genomic data are often highly unequal. Without appropriate frameworks for intellectual property and benefit-sharing, such imbalances may exacerbate global disparities between developed and developing regions. Differences in farmers’ access to data and technology further constrain the equitable implementation of PRS/PGS in real-world agricultural systems.