

open-access GWAS summary statistics and LD reference panels has facilitated reproducibility, cross-platform application, and secondary analyses (Wang et al., 2023).

Across both human medicine and crop breeding, PRS/PGS share a fundamentally analogous objective: to enable early prediction, risk stratification, and selection decisions at the individual level under resource constraints. In medical research, PRS provides a relatively stable estimate of genetic risk across the life course beyond traditional risk factors, supporting stratified screening, longitudinal monitoring, and personalized intervention (Lennon et al., 2024; Xiang et al., 2024). In crop breeding, PGS is methodologically aligned with genomic selection, and is particularly valuable for traits that are costly or late to measure (e.g., perennial crops or complex stress-related traits), where it can serve as an early surrogate phenotype for individual selection, parental optimization, and cross prediction, thereby increasing genetic gain per unit time or cost (Sima et al., 2024). Consequently, establishing a unified methodological language and evaluation framework across medicine and breeding has become an important direction for advancing the application of statistical genetics.

Despite continuous methodological and data advances, the portability and robustness of PRS/PGS across populations remain major challenges. Predictive performance is highly dependent on the allele frequency spectrum and LD structure of the training population. When LD patterns and the tagging relationships between SNPs and causal variants differ across ancestries, signal attenuation or effect size distortion commonly occurs during extrapolation. In addition, population-specific effects, gene-environment interactions, and heterogeneity in phenotype definition and measurement further contribute to reduced explanatory power, calibration bias, and instability of decision thresholds. From a statistical inference perspective, these issues can be understood as an estimand mismatch arising from differences in LD structure, allele frequency spectra, and effect distributions across populations (Duncan et al., 2019; Jayasinghe et al., 2024; Fang, 2026). This structural bias manifests consistently across both human and agricultural systems and extends to concerns regarding population fairness and practical implementation. Moreover, imbalances in sample size and resource availability across ancestries exacerbate the underperformance of PRS in underrepresented populations, making cross-population PRS applications a complex challenge involving statistical modeling, data infrastructure, and ethical governance.

Under this context, it is necessary to re-examine the nature of PRS/PGS from a statistical inference perspective. Strictly speaking, PRS is not a direct estimate of “true genetic risk,” but rather a model-dependent predictive functional defined by the training data, LD structure, and effect estimation model. This perspective is intrinsically consistent with the statistical interpretation of SNP heritability (Fang, 2026): the latter quantifies the proportion of phenotypic variance explained by additive genetic effects under a given set of markers and model assumptions, whereas PRS projects this genetic signal into an individual-level predictive quantity under the same informational constraints. In other words, SNP heritability reflects variance explained at the population level, whereas PRS reflects predictive ability at the individual level, together forming a continuous inference chain from variance decomposition to individual prediction.

Within this unified framework, differences among PRS methods (e.g., C+T, LD-aware Bayesian shrinkage, and multi-ancestry transfer models) fundamentally correspond to different modeling assumptions regarding effect size distributions, LD structure, and sparsity, thereby implying different statistical targets (estimands). Consequently, PRS performance depends not only on sample size and data quality, but also on the degree of alignment between model assumptions and the target population. The decline in cross-population predictive performance can thus be interpreted as a manifestation of estimand mismatch at the level of individual prediction.

Building on this perspective, the present study focuses on the “individual prediction layer” within the broader framework of statistical genetics, extending prior work on methodological paradigm evolution (Fang and Wu, 2026) and variance-based inference (Fang, 2026). We systematically review and compare state-of-the-art PRS/PGS methods in multi-population contexts, including cross-ancestry effect estimation and transfer learning, ancestry-aware LD modeling, functional annotation and causal refinement, as well as model stacking and recalibration strategies, with representative methods such as PRS-CSx, CT-SLEB, and PolyPred. At the level of