

2015; Ni et al., 2018). Furthermore, the SumHer method, based on the LDAK framework, allows SNP effects to depend on minor allele frequency (MAF) and linkage disequilibrium (LD) structure, thereby introducing greater flexibility in modeling genetic architecture (Speed and Balding, 2019).

Despite their shared objective of estimating SNP-based heritability, these methods often yield substantially different results in practice. The UK Biobank (UKB), one of the largest biomedical resources available (Bycroft et al., 2018), provides an ideal setting for systematic comparison. For height—a highly heritable and polygenic trait—typical estimates in UKB European populations show a consistent pattern: GREML-based approaches yield estimates around 0.60~0.69, LDSC produces slightly lower estimates (~0.55-0.60), and SumHer often yields intermediate or slightly higher estimates (~0.63) (Ge et al., 2017; Hou et al., 2019; Speed et al., 2020). Further analyses indicate that, under matched samples and SNP sets, LDSC tends to underestimate heritability by approximately 7%~14% relative to individual-level methods, whereas SumHer may produce estimates that are 5%~38% higher, depending on LD reference panels and model assumptions (Hou et al., 2019; Speed et al., 2020).

These systematic discrepancies raise a fundamental question: are SNP-based heritability estimates obtained from different methods statistically comparable? From a rigorous statistical perspective, the answer is not straightforward. SNP-based heritability is not a fixed biological constant, but rather an estimand—a quantity that depends on the data structure, model specification, and underlying assumptions (Rawlik et al., 2020).

At the data level, GREML leverages individual-level genotype data to construct the GRM, thereby directly capturing genetic similarity between individuals. In contrast, LDSC relies on GWAS summary statistics and external LD reference panels, making its estimates highly sensitive to LD mismatch (Bulik-Sullivan et al., 2015). When LD reference panels do not match the target population, systematic bias may arise (Ni et al., 2018).

At the level of model assumptions, different methods impose distinct constraints on the distribution of genetic effects. Standard GREML assumes homogeneous contributions of SNPs to genetic variance, whereas LDSC adopts a simplified linear model. In contrast, LDAK-based approaches (e.g., SumHer) explicitly allow SNP effects to vary with MAF and LD. Under realistic genetic architectures—where low-frequency variants tend to have larger effects and causal variants are enriched in low-LD regions—such flexible models can substantially increase heritability estimates (Speed et al., 2017).

Linkage disequilibrium and allele frequency distributions play a central role in determining heritability estimates. In real genomes, causal variants are often unevenly distributed, with enrichment in specific regions such as the major histocompatibility complex (MHC). For example, removing the MHC region in UKB analyses can reduce SNP heritability estimates by more than 0.2 for certain traits, highlighting the non-uniform distribution of genetic variance across the genome (Ge et al., 2017). This observation emphasizes that SNP-based heritability reflects the variance that can be captured by observed markers, rather than total genetic variance.

Sample size and statistical efficiency also influence estimation results. In large-scale datasets such as UKB, methods such as randomized Haseman-Elston regression (RHE-reg) and closed-form estimators achieve comparable accuracy to GREML while substantially improving computational efficiency, and further reveal systematic differences between methods (Hou et al., 2019). In addition, participation bias may affect genetic correlations and downstream analyses, but its impact on SNP heritability is generally modest (<5%), indicating relative robustness of variance component estimates (Schoeler et al., 2023).

Taken together, these findings suggest that cross-method differences in SNP heritability do not simply reflect biological variation, but are largely driven by differences in statistical models and data structures. This perspective is particularly important for interpreting the “missing heritability” problem: the gap between SNP-based and pedigree-based estimates is often attributable to incomplete SNP coverage, imperfect LD tagging, and model assumptions, rather than the absence of true genetic effects (Yang et al., 2015).

Based on the above research background, this study takes human height in the UK Biobank database as an entry point and constructs a systematic analytical framework grounded in real data. Within this framework, the focus is