

placed on investigating the statistical origins of discrepancies among different heritability estimation methods, such as GREML, LDSC, and SumHer. Furthermore, the study seeks to distinguish whether these differences arise from the inherent genetic architecture of the trait itself or are introduced by methodological assumptions and model specifications. On this basis, the research attempts to integrate the theoretical foundations and parameter interpretation logic of multiple methods, with the aim of developing a unified statistical interpretation framework that enhances the comparability and consistency of results across approaches. By combining theoretical derivation with empirical analysis, the study more clearly delineates the statistical meaning of SNP heritability and provides a more standardized and robust interpretative paradigm for large-scale genome-wide association studies.

2 Materials and Methods

2.1 Data source: UK biobank cohort

This study is based on data from the UK Biobank (UKB), a large-scale population cohort comprising approximately 500 000 individuals aged 40–69 years, with extensive phenotypic and genome-wide genotypic information (Bycroft et al., 2018).

In this study, data processing and parameter settings were based on established frameworks from large-scale heritability analyses, while also being specifically tailored and optimized according to the characteristics of the research subject. In terms of sample selection, to minimize the potential confounding effects of population structure and relatedness, approximately 290,000 individuals of European ancestry who were unrelated to each other were included. At the level of genetic markers, the study focused on approximately 460,000 common single nucleotide polymorphisms (SNPs), and by applying a minimum allele frequency (MAF) threshold greater than 0.01, effectively excluded the noise introduced by low-frequency variants. For phenotype selection, height was chosen as the trait of interest, as it has high heritability and is jointly regulated by multiple genes, making it a classic model in quantitative genetic research.

These settings are consistent with previous UKB-based heritability analyses and provide statistically stable and high-precision estimates of SNP-based heritability (Ge et al., 2017; Hou et al., 2019). In addition, the large sample size reduces estimation variance and enhances the detectability of systematic differences across methods (Hou et al., 2019).

During the data preprocessing stage, all analyses were conducted under the assumption of stringent quality control. At the individual level, samples with high missingness, discrepancies in reported versus genetic sex, and individuals with heterozygosity rates significantly deviating from the overall distribution were excluded, thereby effectively reducing interference introduced by data anomalies or measurement errors. At the SNP level, further filtering was applied by removing markers with low call rates, those showing significant deviation from Hardy-Weinberg equilibrium, and low-frequency variants, ensuring the reliability and statistical stability of the genetic markers from the outset. Considering that population structure may introduce potential confounding effects on the estimation of genetic parameters, principal component analysis (PCA) was employed to correct for population stratification, thereby mitigating systematic biases arising from differences in genetic background. Collectively, these preprocessing steps establish a robust foundation for data analysis and play a critical role in improving the accuracy and interpretability of heritability estimates (Yang et al., 2010; Bulik-Sullivan et al., 2015).

2.2 Statistical framework for SNP heritability estimation

This study focuses on the estimation of SNP heritability and provides a systematic comparison of three representative methodological approaches. The GREML method, which is based on individual-level genotype data, directly characterizes genetic similarity among individuals by constructing a genomic relationship matrix. In contrast, the LDSC method relies on GWAS summary statistics and, without requiring access to raw individual-level data, decomposes statistical signals through the structure of linkage disequilibrium. Building upon this framework, the SumHer method further introduces more flexible assumptions about genetic architecture by applying weighted modeling to the distribution of effects across loci. These three approaches differ fundamentally in terms of their data requirements, assumptions about the distribution of genetic effects, and the definitions of the