**Glossary**

**Allele frequency:** The proportion of a particular allele in a population's gene pool.

**Clumping:** A statistical method that groups genetic markers in high linkage disequilibrium (LD) together to identify the most significant independent associated loci to reduce redundancy in genome-wide association studies (GWAS) caused by correlation between genetic markers.

**GSMR:** A method that uses summary statistics from GWAS and LD information to estimate causal association between an exposure and outcome variable after detecting and eliminating genetic variants with pleiotropic effect.

**GWAS:** A genome-wide association study that investigates associations between genetic variants and traits or diseases across the entire genome.

**GWAS pairwise:** A method that evaluates the pairwise correlation between genetic variants in GWAS data from two traits to identify potential causal loci that affect both traits.

**HEIDI-outlier:** A statistical method that identifies genetic variant outliers in GWAS data that may be due to genetic heterogeneity or population stratification.

**Inverse variance weighted MR:** A Mendelian randomization method that weights genetic variants by their inverse of variance of their effect on the outcome variables to increase statistical power and accuracy of MR.

**LD:** Linkage disequilibrium, a non-random association between genetic markers on the same chromosome due to the co-inheritance of genetic variants.

**LD-Score regression:** A method that uses GWAS summary statistics and LD information to estimate the heritability and genetic correlation between complex traits.

**Local Analysis of coVariant Annotation (LAVA):** A method that evaluates local heritability and genetic correlation between complex traits in a specific genomic region using GWAS summary statistics.

**Mendelian randomization:** A method that uses genetic variants as instrumental variables to estimate causal effects between exposures and outcomes.

**Minor allele frequency:** The frequency of the least common allele of a genetic marker in a population's gene pool.

**MR-egger:** A method that accounts for directional horizontal pleiotropy when estimating causal effects in Mendelian randomization.

**MR-PRESSO:** A method that identifies and corrects for genetic variant outliers in Mendelian randomization analysis.

**Multivariate Mendelian randomization (MVMR):** A method that estimate causal effects of multiple exposures on outcome variables.

**Pleiotropy:** The phenomenon where a single genetic variant influences multiple traits or diseases.

**Quantitative trait loci:** a statistical method that links trait measurements (e.g. DNA expression) and genotypic data.

**Simple- and weighted-mode:** Mendelian randomization methods that group SNPs based on the similarity of their effect sizes and choose the group with the largest number of SNPs to estimate causal effects.

**SNPs:** Single nucleotide polymorphisms, a type of genetic variant that involves a single nucleotide change in DNA sequence.

**Weighted median:** A method that uses a weighted median estimator to estimate causal effects in Mendelian randomization.