**Suppl 1.** Details on the characteristics of each included dataset

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Phenotype** | **Data source** | **Total sample size** | **Population** | **# SNPs** |
| Depression | Howard, D.M., Adams, M.J., Clarke, TK. et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci 22, 343–352 (2019). | 500,199 (excluding 23andMe) | European | 8.1M |
| Gallstone Disease | Fairfield CJ., Drake TM., Pius R. et al. Genome-wide analysis identifies gallstone-susceptibility loci including genes regulating gastrointestinal motility. Hepatology.75, 1081-1094 (2022). | 28,627 cases  348,373controls | European | 10.1M |
| Smoking, Drinking | Liu M, Jiang Y, Wedow R. et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019 Feb;51(2):237-244. | 1,232,091 | European | 11.8M |
| Education attainment | Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50, 1112-1121(2018). | 1,100,000 | European | 8.1M |

SNP: single nucleotide polymorphism