

MIT Open Access Articles

Phenome-wide heritability analysis of the UK Biobank

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Ge, Tian; Chen, Chia-Yen; Neale, Benjamin M.; Sabuncu, Mert R. and Smoller, Jordan W. "Phenome-Wide Heritability Analysis of the UK Biobank." Edited by Benjamin W. Domingue. PLOS Genetics 13, no. 4 (April 2017): e1006711 © 2017 Ge et al

As Published: <http://dx.doi.org/10.1371/journal.pgen.1006711>

Publisher: Public Library of Science

Persistent URL: <http://hdl.handle.net/1721.1/110042>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution 4.0 International License



RESEARCH ARTICLE

Phenome-wide heritability analysis of the UK Biobank

Tian Ge^{1,2,3*}, Chia-Yen Chen^{2,3,4}, Benjamin M. Neale^{2,3,4}, Mert R. Sabuncu^{1,5}, Jordan W. Smoller^{2,3}✉*

1 Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital / Harvard Medical School, Charlestown, MA, United States of America, **2** Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, United States of America, **3** Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, United States of America, **4** Analytic and Translational Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, United States of America, **5** School of Electrical and Computer Engineering and Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, United States of America

✉ These authors contributed equally to this work.

* tge1@mgh.harvard.edu (TG); jsmoller@hms.harvard.edu (JWS)



OPEN ACCESS

Citation: Ge T, Chen C-Y, Neale BM, Sabuncu MR, Smoller JW (2017) Phenome-wide heritability analysis of the UK Biobank. *PLoS Genet* 13(4): e1006711. <https://doi.org/10.1371/journal.pgen.1006711>

Editor: Benjamin W. Domingue, Stanford University, UNITED STATES

Received: November 22, 2016

Accepted: March 22, 2017

Published: April 7, 2017

Copyright: © 2017 Ge et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The analyses presented in this study were based on data accessed through the UK Biobank <http://www.ukbiobank.ac.uk> on March 3, 2016.

Funding: This research was carried out in part at the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital (MGH), using resources provided by the Center for Functional Neuroimaging Technologies, P41EB015896, a P41 Biotechnology Resource Grant supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB),

Abstract

Heritability estimation provides important information about the relative contribution of genetic and environmental factors to phenotypic variation, and provides an upper bound for the utility of genetic risk prediction models. Recent technological and statistical advances have enabled the estimation of additive heritability attributable to common genetic variants (SNP heritability) across a broad phenotypic spectrum. Here, we present a computationally and memory efficient heritability estimation method that can handle large sample sizes, and report the SNP heritability for 551 complex traits derived from the interim data release (152,736 subjects) of the large-scale, population-based UK Biobank, comprising both quantitative phenotypes and disease codes. We demonstrate that common genetic variation contributes to a broad array of quantitative traits and human diseases in the UK population, and identify phenotypes whose heritability is moderated by age (e.g., a majority of physical measures including height and body mass index), sex (e.g., blood pressure related traits) and socioeconomic status (education). Our study represents the first comprehensive phenome-wide heritability analysis in the UK Biobank, and underscores the importance of considering population characteristics in interpreting heritability.

Author summary

Heritability of a trait refers to the proportion of phenotypic variation that is due to genetic variation among individuals. It provides important information about the genetic basis of complex traits and indicates whether a phenotype is an appropriate target for more specific statistical and molecular genetic analyses. Recent studies have leveraged the increasingly ubiquitous genome-wide data and documented the heritability attributable to common genetic variation captured by genotyping microarrays for a wide range of human traits. However, heritability is not a fixed property of a phenotype and can vary with population-

National Institutes of Health (NIH). This work also involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program; specifically, grant numbers S10RR023043 and S10RR023401. This research was also funded in part by NIH grants R01NS083534, R01NS070963, 1K25EB013649-01 and 1R21AG050122-01A1 (to MRS); K24MH094614 (to JWS); and an MGH ECOR Tosteson Postdoctoral Fellowship Award (to TG). JWS is a Tepper Family MGH Research Scholar and was also supported in part by a gift from the Demarest Lloyd, Jr. Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This research has been conducted using the UK Biobank resource.

Competing interests: The authors have declared that no competing interests exist.

specific differences in the genetic background and environmental variation. Here, using a computationally and memory efficient heritability estimation method, we report the heritability for a large number of traits derived from the large-scale, population-based UK Biobank, and, for the first time, demonstrate the moderating effect of three major demographic variables (age, sex and socioeconomic status) on heritability estimates derived from genome-wide common genetic variation. Our study represents the first comprehensive heritability analysis across the phenotypic spectrum in the UK Biobank.

Introduction

The heritability of a trait refers to the proportion of phenotypic variance that is attributable to genetic variation among individuals. Heritability is commonly measured as either the contribution of total genetic variation (broad-sense heritability, H^2), or the fraction due to additive genetic variation (narrow-sense heritability, h^2) [1]. A large body of evidence from twin studies has documented that essentially all human complex traits are heritable. For example, a recent meta-analysis of virtually all twin studies published between 1958 and 2012, encompassing 17,804 traits, reported that the overall narrow-sense heritability estimate across all human traits was 49%, although estimates varied widely across phenotypic domains [2]. Over the past decade, the availability of genome-wide genotyping has enabled the direct estimation of additive heritability attributable to common genetic variation (“SNP heritability” or h^2_{SNP}) [3–5]. These estimates do not capture non-additive genetic effects such as dominance or epistasis, and provide a lower bound for narrow-sense heritability because they also do not capture contributions (e.g., from rare variants) that are not assayed by most genotyping microarrays and are not well tagged by genotyped variants. Nevertheless, estimates of SNP heritability can provide important information about the genetic basis of complex traits such as the proportion of phenotypic variation that could be explained by common-variant genome-wide association studies (GWAS).

However, heritability is not a fixed property of a phenotype but depends on the population in which it is estimated. As a ratio of variances, it can vary with population-specific differences in both genetic background and environmental variation [1]. For example, twin data have documented variations in the heritability of childhood IQ by socioeconomic status (SES) [6], highlighting that different environment may have different relative contributions to the variance of a phenotype. In addition, heritability estimates for a range of complex phenotypes have been shown to vary according to the sex and age distributions of the sampled populations [2]. Identifying variables that may affect the heritability of complex traits has implications for the design of GWAS, highlighting subgroups and environmental conditions in which common-variant contributions may be diminished or magnified. To date, however, studies of complex trait heritability and the effect of modifying variables have produced mixed results likely due to sample size limitations and population-specific differences in genetic and environmental variance that may be operating in different cohorts.

The UK Biobank (<http://www.ukbiobank.ac.uk>) provides a unique opportunity to estimate the heritability of traits across a broad phenotypic spectrum in a single population sample. The UK Biobank is a large prospective population-based cohort study that enrolled 500,000 participants aged 40–69 years between 2006 and 2010 [7]. The study has collected a wealth of phenotypic data from questionnaires, physical and biological measurements, and electronic health records as well as genome-wide genotype data. However, this rich data source also presents analytic challenges. For example, with the large sample size, existing heritability estimation

methods such as genome-wide complex trait analysis (GCTA) [3–5] and LD (linkage disequilibrium) score regression [8] become computationally expensive and memory intensive, and thus can be difficult to apply. Here we implemented a computationally and memory efficient approach to estimate the heritability for 551 complex traits derived from the interim data release (152,736 subjects) of the UK Biobank, comprising both quantitative phenotypes and disease categories. We then examined how heritability estimates are modified by three major demographic variables: age, sex and socioeconomic status (SES). Our results underscore the importance of considering population characteristics in estimating and interpreting heritability, and may inform efforts to apply genetic risk prediction models for a broad range of human phenotypes.

Results

We report the heritability for 551 traits that were made available to us through the interim data release of the UK Biobank (downloaded on Mar 3, 2016) and had sufficient sample sizes to achieve accurate heritability estimates (standard error of the heritability estimate smaller than 0.1; 15 disease codes excluded) using a computationally and memory efficient heritability estimation method (see [Methods](#) and [S1 Text](#)).

The 551 traits can be classified into 11 general phenotypic domains as defined by the UK Biobank to group individual data fields into broadly related sets: cognitive function (5 traits), early life factors (7 traits), health and medical history (60 traits), hospital in-patient main diagnosis ICD-10 codes (194 traits), life style and environment (88 traits), physical measures (50 traits), psychosocial factors (40 traits), self-reported cancer codes (9 traits), self-reported non-cancer illness codes (79 traits), sex-specific factors (14 traits), and sociodemographics (5 traits). ICD-10 (the International Classification of Diseases, version-10) is a medical classification list published by the World Health Organization (WHO), which contains thousands of diagnostic codes. [Fig 1](#) shows the percentage of each domain that makes up the 551 traits we analyzed. Using the top-level categories and chapters of the self-reported disease and ICD-10 coding tree, we can further break down self-reported non-cancer illness codes and ICD-10 codes into different functional domains ([S1 Fig](#)). We note that since we only analyzed disease codes that had prevalence greater than 1% in the sample, distribution of the disease traits across functional domains was skewed. For example, we investigated a large number of gastrointestinal and musculoskeletal traits, while diseases that have low prevalence in the sampled population such as psychiatric disorders were not well represented.

[Table 1](#) lists the top heritable traits in each domain (the most heritable trait and traits with heritability estimates greater than 0.30). [S1 Table](#) and [S2 Table](#) show the heritability estimates, standard error estimates, sample sizes, covariates adjusted, prevalence in the sample (for binary traits) and other relevant information for all the traits we analyzed. Common genetic variants appear to have an influence on most traits we investigated, although heritability estimates showed heterogeneity within and across trait domains. Complex traits that exhibited high SNP heritability (larger than 0.40) included human height (0.685±0.004), skin color (very fair/fair vs. other, 0.556±0.008), ease of skin tanning (very/moderately tanned vs. mildly/occasionally/never tanned, 0.454±0.006), comparative height at age 10 (taller than average, 0.439±0.007; shorter than average, 0.405±0.008), rheumatoid arthritis (0.821±0.046), hypothyroidism/myxedema (0.814±0.017), malignant neoplasm of prostate (0.426±0.093), and diabetes diagnosed by doctor (0.414±0.016), among others. On the other end of the spectrum, traits such as duration of walks/moderate activity/vigorous activity, frequency of stair climbing, ever had stillbirth, spontaneous miscarriage or termination, painful gums, stomach disorder, fracture, injuries to the head/knee/leg, and pain in joint had zero or close to

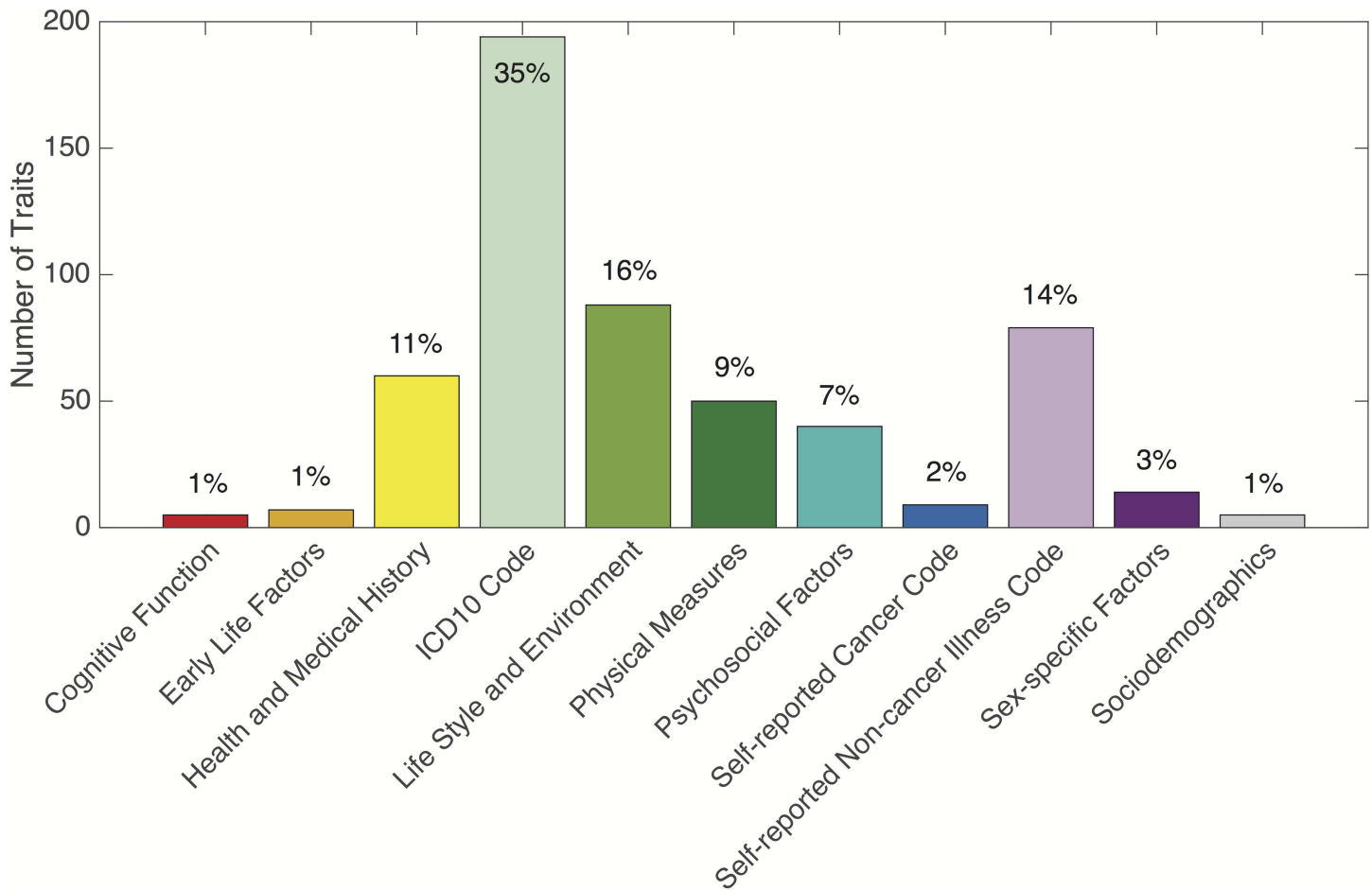


Fig 1. The number of traits in each of the 11 phenotypic domains that make up the 551 traits analyzed in the UK Biobank: cognitive function (5 traits), early life factors (7 traits), health and medical history (60 traits), hospital in-patient main diagnosis ICD-10 codes (194 traits), life style and environment (88 traits), physical measures (50 traits), psychosocial factors (40 traits), self-reported cancer codes (9 traits), self-reported non-cancer illness codes (79 traits), sex-specific factors (14 traits), and sociodemographics (5 traits).

<https://doi.org/10.1371/journal.pgen.1006711.g001>

zero heritability estimates, indicating that their phenotypic variation is largely determined by environmental factors, or there is widespread heterogeneity or substantial measurement error in these phenotypes. SNP heritability estimates for several phenotypes, including diseases with known immune-mediated pathogenesis (rheumatoid arthritis, psoriasis, diabetes, hypothyroidism), were markedly reduced when the major histocompatibility complex (MHC) region was excluded from analysis (S4 Table), and thus need to be interpreted with caution (see Discussion).

A substantial fraction of the phenotypes we examined were based on self-report illness codes or diagnostic (ICD-10) codes, which may be noisy and have low specificity. However, the SNP heritability estimates for 14 pairs of self-reported illness and ICD-10 codes that represent the same or closely matched diseases were largely consistent and had a Pearson correlation of 0.78 (Table 2), indicating that both phenotypic approaches captured useful and comparable variations in these phenotypes.

Heritability analysis stratified by sex identified a number of traits whose heritability showed significant difference in males and females after multiple testing correction (Fig 2). For example, the analyses of diastolic and systolic blood pressure, and self-reported hypertension and

Table 1. The SNP heritability estimates (h²), standard error (SE) estimates, sample sizes (N), and prevalence (prev) in the sample (for binary traits) for the top heritable traits in each phenotypic domain. All heritability estimates adjusted for genotyping array, UK Biobank assessment center, age at recruitment, sex (except male-specific traits indicated by *) and top 10 principal components of the genotype data as covariates. [‡]: additionally adjusted for weight; [§]: additionally adjusted for negative experience. See S1 Table and S2 Table for the way we binarized categorical variables and other relevant information.

Field ID	Field Name	Type	N	Prev	h ²	SE
20016	Fluid intelligence score	continuous	34491	NA	0.233	0.011
1697	Comparative height at age 10	categorical	106497	25.36%	0.439	0.007
2976	Age diabetes diagnosed	continuous	5369	NA	1.000	0.074
6148	Eye problems/disorders	categorical	33750	2.04%	0.722	0.097
2443	Diabetes diagnosed by doctor	binary	107935	5.25%	0.414	0.016
6152	Blood clot, DVT, bronchitis, emphysema, asthma, rhinitis, eczema, allergy diagnosed by doctor	categorical	108022	12.37%	0.344	0.010
41202	C61 Malignant neoplasm of prostate	binary	32780	2.24%	0.426*	0.093
41202	C60-C63 Malignant neoplasms of male genital organs	binary	32780	2.46%	0.322*	0.087
1747	Hair colour (natural, before greying)	categorical	107947	4.62%	1.000	0.018
1717	Skin colour	categorical	106773	80.09%	0.556	0.008
1727	Ease of skin tanning	categorical	105966	61.13%	0.454	0.006
6144	Never eat eggs, dairy, wheat, sugar	categorical	107832	2.68%	0.691	0.026
50	Standing height	continuous	107976	NA	0.685	0.004
23105	Basal metabolic rate	continuous	106311	NA	0.397	0.004
23102	Whole body water mass	continuous	106252	NA	0.331 [‡]	0.004
23130	Trunk predicted mass	continuous	106157	NA	0.325 [‡]	0.004
23106	Impedance of whole body	continuous	106304	NA	0.307	0.004
20151	Forced vital capacity (FVC), best measure	continuous	84301	NA	0.339	0.005
20150	Forced expiratory volume in 1-second (FEV1), best measure	continuous	84305	NA	0.324	0.005
3148	Heel bone mineral density (BMD)	continuous	62546	NA	0.327	0.006
20126	Bipolar and major depression status	categorical	25571	19.91%	0.195 [§]	0.032
20001	Prostate cancer	binary	50997	1.59%	0.310*	0.077
20002	Psoriasis	binary	108148	1.15%	1.000	0.047
20002	Thyroid problem (not cancer)	binary	108148	5.77%	0.876	0.015
20002	Rheumatoid arthritis	binary	108148	1.18%	0.821	0.046
20002	Hypothyroidism/Myxoedema	binary	108148	4.83%	0.814	0.017
20002	Diabetes	binary	108148	5.14%	0.396	0.016
20002	Dermatology	binary	108148	4.27%	0.368	0.019
20002	Asthma	binary	108148	12.44%	0.340	0.010
2395	Hair/Balding pattern	categorical	50660	18.51%	0.479*	0.017
2375	Relative age of first facial hair	categorical	49384	6.64%	0.303*	0.030
6138	Qualifications	categorical	107158	30.16%	0.294	0.007

<https://doi.org/10.1371/journal.pgen.1006711.t001>

high blood pressure provided consistent evidence that the heritability of blood pressure related traits and diseases is significantly higher in females than in males.

A majority of physical measures showed decreasing heritability with age (S3 Table). More specifically, 33 out of 50 physical measures had a significant decreasing trend in heritability estimates after accounting for multiple testing correction (mean slope of the 33 traits -0.0035, i.e., heritability estimates decrease by 3.5 percent per decade). The age-varying SNP heritability estimates and their standard errors for 12 traits that showed both significant slopes and significantly different heritability estimates between the first (40–49 years) and last age range (64–73 years) are shown in Fig 3A. S2 Fig shows the mean and standard deviation of the 12 traits in each age range.

Table 2. A head-to-head comparison of SNP heritability estimates (h2) for the self-reported illness codes and ICD-10 codes that represent the same or closely matched diseases.

Self-reported disease	h2 (self-reported)	ICD-10 code	h2 (ICD-10)
Breast cancer	0.133	C50 Malignant neoplasm of breast	0.207
Skin cancer	0.061	C43-C44 Melanoma and other malignant neoplasms of skin	0.160
Male genital tract cancer	0.226	C60-C63 Malignant neoplasms of male genital organs	0.322
Prostate cancer	0.310	C61 Malignant neoplasm of prostate	0.426
Cerebrovascular disease	0.066	I60-I69 Cerebrovascular diseases	0.197
Angina	0.176	I20 Angina pectoris	0.136
Liver/Biliary/Pancreas problem	0.074	K80-K87 Disorders of gallbladder, biliary tract and pancreas	0.099
Heart attack/Myocardial infarction	0.184	I21 Acute myocardial infarction	0.193
Gastro-oesophageal reflux (gord)/Gastric reflux	0.099	K21 Gastro-oesophageal reflux disease	0.105
Cholelithiasis/Gall stones	0.123	K80 Cholelithiasis	0.143
Uterine fibroids	0.087	D25 Leiomyoma of uterus	0.054
Enlarged prostate	0.157	N40 Hyperplasia of prostate	0.127
Pneumonia	0.075	J18 Pneumonia—organism unspecified	0.160
Diverticular disease/Diverticulitis	0.195	K57 Diverticular disease of intestine	0.179

<https://doi.org/10.1371/journal.pgen.1006711.t002>

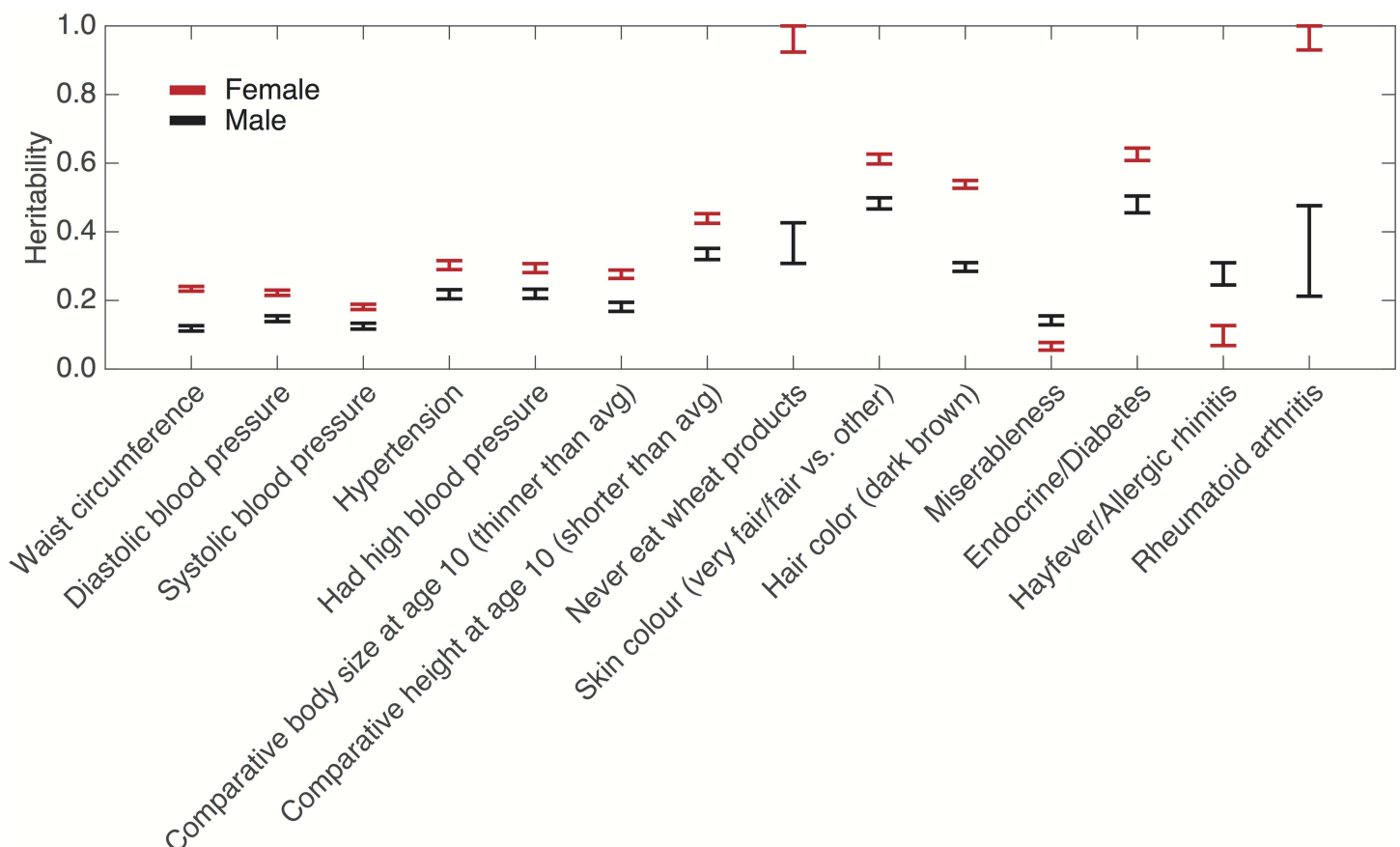


Fig 2. SNP heritability estimates and their standard errors of the traits in the UK Biobank that show significantly different SNP heritability in females and males after multiple testing correction. The heritability estimates of rheumatoid arthritis, endocrine/diabetes and wheat products intake reported here are based on genome-wide SNPs and will be markedly reduced when the major histocompatibility complex (MHC) region (chr6:25-35Mb) is excluded from analysis, and thus need to be interpreted with caution.

<https://doi.org/10.1371/journal.pgen.1006711.g002>

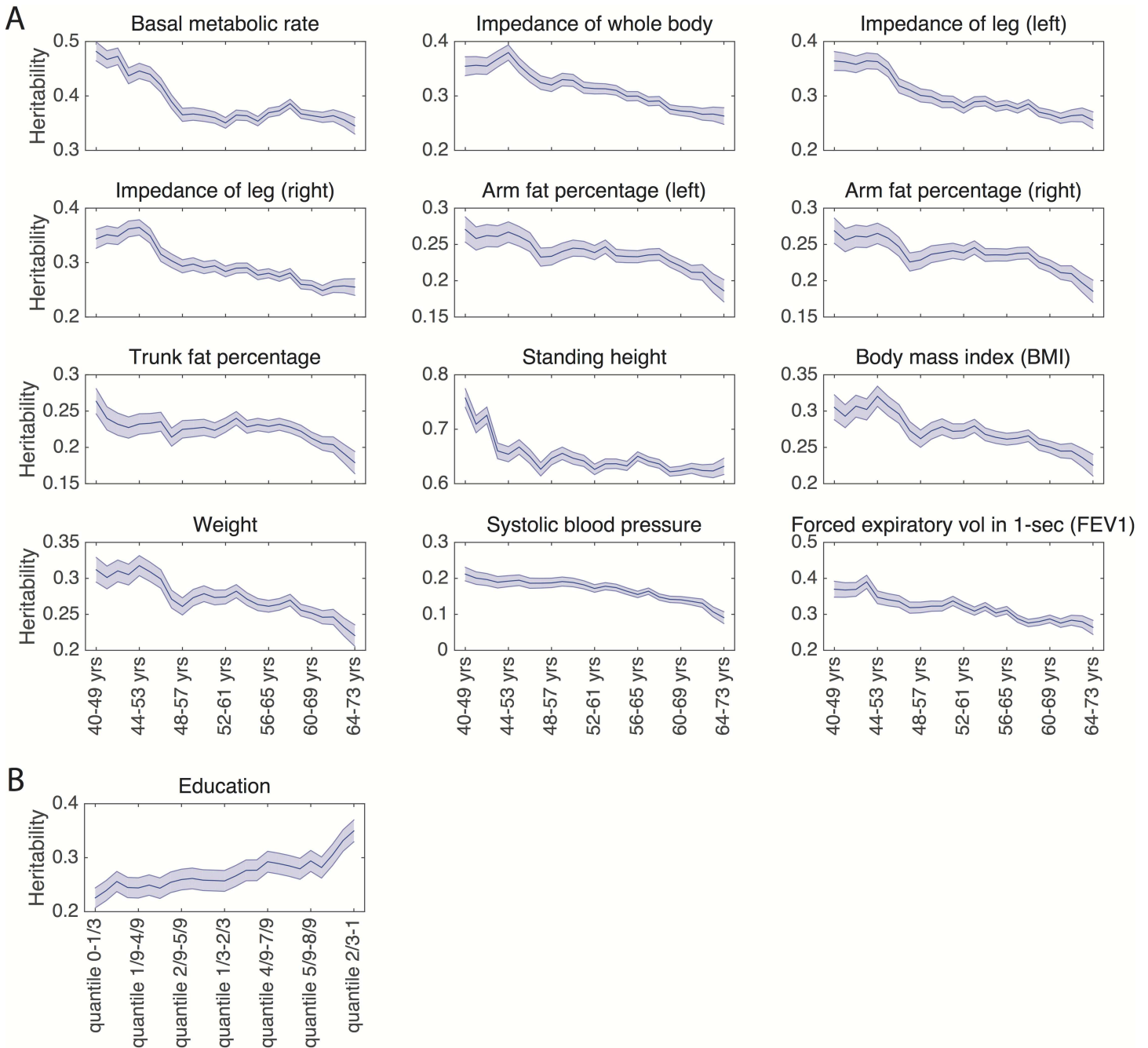


Fig 3. (A) The age-varying SNP heritability estimates and their standard errors (shaded region) for the 12 traits whose heritability significantly decreases with age after accounting for multiple testing correction; (B) The stratified heritability estimates and standard errors (shaded region) of education (has college or university degree or not), on which the socioeconomic status (SES) measured by the Townsend deprivation index has a significant moderating effect after multiple testing correction.

<https://doi.org/10.1371/journal.pgen.1006711.g003>

When we stratified heritability by the Townsend deprivation index, a proxy for SES, education (has college or university degree or not) was the only trait on which SES had a significant moderating effect after accounting for multiple testing correction. Fig 3B shows that the heritability of education increases with increasing SES.

Discussion

Estimating the heritability of complex, polygenic traits is an important component of defining the genetic basis of human phenotypes. In addition, heritability estimates provide a theoretical upper bound for the utility of genetic risk prediction models [9]. We calculated the common-variant heritability of 551 phenotypes derived from the interim data release of the UK Biobank, and confirmed that common genetic variation contributes to a broad array of quantitative traits and human diseases in the UK population. Two aspects of our work are particularly notable. First, we developed a computationally and memory efficient method that enabled us to calculate the most extensive population-based survey of SNP heritability to date. Second, we found that the heritability for a number of phenotypes is moderated by major demographic variables, demonstrating the dependence of heritability on population characteristics. We discuss each of these advances and the limitations of the biobank data and our analyses below.

Classical methods to estimate SNP heritability, such as GCTA (also known as the GREML method), rely on the restricted maximum likelihood (ReML) algorithm [3–5], which can give unbiased heritability estimates in quantitative trait analysis and non-ascertained case-control studies, and is statistically efficient when the trait is Gaussian distributed [10]. However, ReML is an iterative optimization algorithm, which is computationally and memory intensive, and thus can be difficult to apply when analyzing data sets with hundreds of thousands of subjects. An alternative and widely used SNP heritability estimation method is LD score regression, which is based on GWAS summary statistics and an external reference panel for the LD structure [8]. The approach can thus be easily applied to complex traits on which large-scale GWAS results are available, and allows meta-analysis of heritability estimates from different studies. Recently, LD score regression has been extended to partition heritability by functional annotation [11], and to estimate the genetic correlation between two traits [12,13]. However, when applying LD score regression to novel phenotypes in a large cohort, conducting GWAS is often time-consuming.

In the present study, we implemented a computationally and memory efficient moment-matching method for heritability estimation, which is closely related to the Haseman-Elston regression [14–16] and phenotype-correlation genetic-correlation (PCGC) regression [10], and produces unbiased SNP heritability estimates for both continuous and binary traits. The moment-matching method is theoretically less statistically efficient than the ReML algorithm (i.e., produces larger standard error on the point estimate) when analyzing quantitative traits, but the power loss is expected to be small [17] and is less of an issue given large sample sizes, such as in the UK Biobank. The moment-matching method is also mathematically equivalent to LD score regression if the following conditions are satisfied: (1) the out-of-sample LD scores estimated from the reference panel and the in-sample LD scores estimated from individual-level genotype data are identical; (2) the intercept in the LD score regression model is constrained to 1 (i.e., assuming that there is no confound and population stratification in the data); and (3) a particular weight is used in the LD score regression (more specifically, the reciprocal of the LD score, which is close to the default setting in the LD score regression software) [18]. Here, since we have constrained our analysis to a white British (Caucasian) sample and have accounted for potential population stratification by including top PCs of the genotype data as covariates, the two methods should produce similar estimates. See [Box 1](#) for an empirical comparison between the moment-matching method, LD score regression and GCTA.

Using the moment-matching method, we found that a large number of traits we examined display significant heritability. For traits whose heritability has been intensively studied, our estimates are generally in line with prior studies. For example, twin and pedigree studies have

Box 1. An empirical comparison between the moment-matching method, LD score regression and genome-wide complex trait analysis (GCTA). To confirm our theoretical expectation of how the moment-matching method compares to LD score regression and GCTA, the most widely used SNP heritability estimation methods, we performed the following empirical analysis. For each trait in [S1 Table](#), we randomly sampled 10,000 subjects and compared SNP heritability estimates from the moment-matching method, LD score regression (with constrained intercept and the regression weight set to the reciprocal of the LD score), and GCTA. [S3 Fig](#) shows that the moment-matching method produced virtually identical estimates to LD score regression, and consistent heritability and standard errors estimates with GCTA. The power loss of the moment-matching method relative to GCTA is negligible. Although the three methods produced consistent heritability estimates, the moment-matching method has much lower computational demand than the ReML algorithm and LD score regression. Specifically, using a single core of the (dual CPU) Intel Xeon 5472 3.0GHz processor and 7 Gb of virtual memory, GCTA takes approximately half an hour to analyze one trait (assuming that the genetic similarity matrix has been pre-computed) and each GWAS takes several hours, while the moment-matching method can produce the SNP heritability estimate of a trait in less than half a minute. When analyzing the full sample (~100,000 subjects), GCTA becomes very difficult to apply because its memory demand increases quadratically and computational complexity scales cubically with the sample size. It is also challenging to conduct stratification analyses across the phenotypic spectrum using LD score regression due to the large number of GWAS to be performed. In contrast, the moment-matching method only takes approximately 20 mins to complete the main heritability analysis and all the stratification analyses for a trait with 100,000 subjects using one CPU processor and 7Gb of memory. To summarize, we have used a method that balances statistical efficiency and computational burden with the flexibility to achieve our main goal in this study.

estimated the heritability of human height and body mass index (BMI) to be approximately 80% and 40–60% [see e.g., 19–21], respectively, although recent studies have shown that heritability may be overestimated in family studies due to, for instance, improper modeling of common environment, assortative mating in humans, genetic interactions, and suboptimal statistical methods [10,22–25]. Using genome-wide SNP data from unrelated individuals, it has been shown that common SNPs explain a large proportion of the height and BMI variation in the population, although SNP heritability estimates are lower than twin estimates [4,5,26]. Specifically, the first GCTA analysis estimated the SNP heritability of human height to be 0.45 using relatively sparse genotyping data (approximately 300,000 SNPs) and showed that the estimate could be higher if imperfect LD between SNPs and causal variants are corrected [4]. A more recent study leveraging whole-genome sequencing data and imputed genetic variants concluded that narrow-sense heritability is likely to be 60–70% for height and 30–40% for BMI [27]. Here, we estimated the SNP heritability of human height and BMI to be 0.685±/−0.004 and 0.274±/−0.004, respectively, which are comparable to the expected range. The SNP heritability estimates of other complex traits of interest, such as age at menarche in girls (0.239±/−0.007), diastolic (0.184±/−0.004) and systolic (0.156±/−0.004) blood pressures, education (has colleague or university degree or not, 0.294±/−0.007), neuroticism (0.130±/−0.005), smoking (ever smoked or not, 0.174±/−0.006), asthma (0.340±/−0.010) and hypertension (0.263±/−0.007) were also more modest and lower than twin estimates, as expected [2].

Heritability is, by definition, a ratio of variances, reflecting the proportion of phenotypic variance attributable to individual differences in genotypes. Because the genetic architecture and non-genetic influences on a trait may differ depending on the population sampled, heritability itself may vary. Examples of this have been reported in the twin literature. In one well-known study, Turkheimer and colleagues [6] reported that the heritability of IQ is moderated by SES in a sample of 320 7-year-old twin pairs of mixed ancestry. In that study, the heritability of IQ was essentially 0 at the lowest end of SES but substantial at the highest end. Subsequent studies of the moderating effects of SES on the heritability of cognitive ability and development using twin designs have produced mixed results [28–33]. In our analysis, using SNP data, we observed no moderating effect of SES (as measured by the Townsend deprivation index) on the heritability of cognitive traits (including fluid intelligence), possibly due to the age range of participants in the UK Biobank (middle and old age) in contrast to many previous studies targeting childhood or early adulthood, and the cross-national differences in gene-by-SES interaction on intelligence as shown by a recent meta-analysis [34]. In addition, the brief cognitive tests available in the UK Biobank may have had limited sensitivity for capturing individual differences in IQ (see [discussion](#) below). On the other hand, the heritability of education showed significant interactions with SES, with increasing heritability at higher SES levels. Prior evidence has suggested that education has substantial genetic correlation with IQ and may be a suitable proxy phenotype for genetic analyses of cognitive performance [35]; thus our results may indirectly support earlier studies of the SES moderation of IQ heritability.

With two exceptions, significant sex differences we observed indicated greater heritability for women compared to men. Our results are consistent with findings from some twin studies but not others. For example, we found that women exhibited significantly greater heritability for measured waist circumference and blood pressure. Twin studies have also reported greater female heritability for waist circumference [36] but no substantial sex difference in heritability of blood pressure [2,37]. A substantial difference between the heritability of rheumatoid arthritis (RA) in males compared to females was observed, although the MHC region has a large impact on the SNP heritability estimates of autoimmune diseases, and thus this finding needs to be interpreted with caution (see [discussion](#) below). While RA is known to be more common in women, a twin analysis found no sex difference in heritability among Finnish and UK twin pairs, though power was limited in that analysis [38]. Intriguingly, greater heritability was observed among men for the personality trait of miserableness, a component of neuroticism, suggesting that environmental factors may be more influential for this trait among women or that measurement error differs by sex.

We examined age effects on heritability for a subset of variables and found that a number of physical measures indexing body size, adiposity, height, as well as systolic blood pressure and lung function, showed declining heritability with age. Age-related declines in heritability may reflect the cumulative effect of environmental perturbations over the lifespan. Prior twin studies of age effects on the heritability of anthropometric traits in adults have had inconsistent results [39–41]. Haworth and colleagues showed that the heritability of BMI increases over childhood [42]. A recent meta-analysis of 32 twin studies documented a non-monotonic relationship between BMI heritability and age (from childhood to late adulthood), with a peak around age 20 and decline thereafter [43]. An age-related decline in indices of body size may reflect a decreasing contribution of genetically-regulated growth processes over the lifespan. However, we were unable to assess the entire trajectory of heritability due to the age range (40–73 years) of the UK Biobank participants. Some but not all studies have also suggested varying or declining heritability with age for blood pressure, lung function and age at first birth [39,44–50].

Our results should be interpreted in light of the limitations associated with the biobank data. First, the UK Biobank is restricted to middle and old age groups, which may be subject to sample selection bias. For example, older and physically/cognitively impaired subjects may be underrepresented in the study, which may have an impact on the heritability estimates stratified by age. Mortality selection can also alter the results of genetic analyses as shown by recent analyses [51]. In addition, the UK Biobank participants comprised a relatively high proportion of well-educated, skilled professionals [52], potentially leading to the underrepresentation or restricted range of certain traits such as smoking relative to other cohorts. Therefore, our heritability estimates may be specific to this UK population and may not generalize to other settings or ancestry groups. Second, although the UK Biobank has collected a wealth of phenotypes, measurements associated with a particular phenotypic domain may not be comprehensive. For example, only five cognitive tests were included in the UK Biobank. The reasoning task (fluid intelligence test) was brief and had a narrow range; the reaction time was averaged from a small number of trials; and the visual memory test (pairs-matching test) had a significant floor effect (a large number of participants made zero or very few mistakes, and thus the scores do not fully reflect individual differences). In addition, all cognitive tests had relatively low reliability across repeat measurements [53]. These noisy measurements may thus downwardly bias heritability estimates of cognition. The Townsend deprivation index, which we used to stratify phenotypes, was calculated based on the national census output area of each participant in which their postcode was located at the time of recruitment, and thus can only serve as a proxy for SES. Third, the phenotypes were limited to those for which we had sufficient data to estimate heritability with adequate precision. Therefore, diseases with low prevalence in the sampled population were not well represented in our analysis. We expect to analyze traits with lower prevalence (e.g., 0.5%) when the genetic data for all UK Biobank participants becomes available. We also assumed in our analysis that the population prevalence of a binary trait is identical to the observed sample prevalence, but diseases such as schizophrenia and stroke are naturally under-ascertained and thus their sample prevalence is often lower than population prevalence. In addition, we note that since we used medical history to define cases and controls, the prevalence of many diseases we investigated reflected lifetime prevalence, which may be different from cross-sectional prevalence used in other studies. We also binarized categorical (multinomial or ordinal) variables to facilitate analysis, but this might not optimally represent variation in these variables with respect to heritability. Fourth, a substantial fraction of the phenotypes we examined were based on self-report or diagnostic (ICD-10) codes, which may or may not validly capture the phenotypes they represent. For example, a recent UK Biobank study shows that 51% of the participants who reported RA were not on RA-relevant medication, a proxy measure of valid diagnosis [54]. However, our head-to-head comparison of the heritability estimates between self-reported illness and ICD-10 codes showed largely consistent results, indicating that both phenotypic approaches at least captured comparable variations in these phenotypes. Prior research evaluating phenotypes derived from electronic health records (EHR) indicate that greater phenotypic validity can be achieved when diagnostic codes are supplemented with text mining methods [55–58]. The specificity of the disease codes might also be improved by leveraging the medication records in the UK Biobank.

Methodologically, our SNP heritability estimation approach, despite its superior computational and memory performance compared to existing methods, also has several limitations. First, heritability estimation always relies on a number of assumptions on the genetic architecture. For example, the moment-matching method we used here, as well as the established GCTA and LD score regression approaches, implicitly assumes that the causal SNPs are randomly spread over the genome, which is independent of the MAF spectrum and the LD

structure, and the effect sizes of causal SNPs are Gaussian distributed and have a specific relationship to their MAFs. Although it has been shown that SNP heritability estimates are reasonably robust to many of these modeling assumptions [59], the estimates can be biased if, for instance, causal SNPs are rarer or more common than uniformly distributed on the MAF spectrum, or are enriched in high or low LD regions across the genome. For example, the heritability estimates for some autoimmune diseases such as psoriasis and RA dropped dramatically when the MHC region (chr6:25-35Mb) was removed when constructing the genetic similarity matrix, indicating, as expected, that causal variants for these diseases are disproportionately enriched in the MHC region. [S4 Table](#) lists all the traits whose heritability estimates decreased by 0.2 or more when the MHC region was taken out, and thus need to be interpreted with caution. Methods to correct for MAF properties and region-specific LD heterogeneity of causal variants have been proposed [27,59,60]. For example, we can stratify MAF and LD structure into different bins, compute a genetic similarity matrix within each bin, and fit a mixed effects model with multiple variance components [27,60]. This approach can give heritability estimates that are more robust to properties of the underlying genetic architecture, but has the downside of increased computational burden and reduced statistical power. A different direction to explore is to estimate SNP heritability using imputed data (in contrast to the genotype data here), which might capture more genetic variation from rare variants, or common variants that are not well tagged by the genotyped SNPs, and thus lead to increased heritability estimates. Second, heritability analysis models, including the one we employed in the present study, typically assume that genetic and environmental effects are independent, i.e., no gene-by-environment (GxE) interaction exists. This is certainly a simplification of the real world where GxE interactions are expected for many complex traits. Recent computational studies have also shown that ignoring GxE interactions in heritability analysis can produce biased estimates [25]. However, modeling GxE would require collecting relevant environmental variables for each phenotype and more sophisticated statistical modeling approaches, e.g., incorporating multiple random effects in the heritability analysis model [5,61]. Due to the limited measurements of environment collected by the UK Biobank, and the extensive analyses we have conducted across the phenotypic spectrum, explicitly modeling the environmental factors and GxE interactions is not feasible. We therefore took an alternative approach to examine the moderating effects of three major demographic variables on heritability estimates by stratifying samples. Of note, consistent heritability estimates across different levels of the stratifying variable do not completely eliminate the potential existence of GxE interactions. Specifically, recent studies have identified genetic heterogeneity in human traits such as BMI and fertility [62,63], indicating that the genetic architecture of a trait may be different across environments (i.e., the genetic correlation of a trait in different environments may be significantly smaller than 1) even if the overall heritability estimates are similar. Dissection of common and unique environmental influences and their interactive effects with genetics on different complex traits, and the shared and unique genetic effects across environments are important future directions to explore. Lastly, as reviewed in [64], a number of empirical genetic similarity measurements computable from genome-wide SNP data have been proposed, which, when utilized in heritability analysis, can give different estimates with different interpretations. In addition, recent studies have argued that estimation error associated with genetic similarity measurements and the ill-posedness of the empirical genetic similarity matrix may produce unstable and unreliable SNP heritability estimates [65]. However, this is an area under active investigation and debate [64–67]. Here, as the first study to screen all UK Biobank variables and provide an overview of the distribution of SNP heritability across different trait domains, and to examine the effect of potential modifying variables on heritability estimates, we used a straightforward and classical modeling approach that is most widely used. To obtain more insights into the genetic

architecture and find the most appropriate and robust model for each individual trait, more systematic investigation is needed.

In sum, using a computationally and memory efficient approach, we provide estimates of the SNP heritability for 551 complex traits across the phenome captured in the population-based UK Biobank. We further identify phenotypes for which the contribution of genetic variation is modified by demographic factors. These results underscore the importance of considering population characteristics in interpreting heritability, highlight phenotypes and subgroups that may warrant priority for genetic association studies, and may inform efforts to apply genetic risk prediction models for a broad range of human phenotypes.

Materials and methods

Ethics statement

This study utilized deidentified data from the baseline assessment of the UK Biobank, a prospective cohort study of 500,000 individuals (age 40–69 years) recruited across Great Britain during 2006–2010 [7]. The protocol and consent were approved by the UK Biobank's Research Ethics Committee.

Participants and data sources

The UK Biobank collected phenotypic data from a variety of sources including questionnaires regarding mental and physical health, food intake, family history and lifestyle, a baseline physical assessment, computerized cognitive testing, linkage with health records, and blood samples for biochemical and DNA analysis. Details about the UK Biobank project are provided at <http://www.ukbiobank.ac.uk>. Data for the current analyses were obtained under an approved data request (Ref: 13905).

Genotyping and quality control

The interim release of the genotype data for the UK Biobank (downloaded on Mar 3, 2016) comprises 152,736 samples. Two closely related arrays from Affymetrix, the UK BiLEVE Axiom array and the UK Biobank Axiom array, were used to genotype approximately 800,000 markers with good genome-wide coverage. Details of the design of the arrays and sample processing can be found at <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=146640> and <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=155583>.

Prior to the release of the genotype data, stringent quality control (QC) was performed at the Wellcome Trust Centre for Human Genetics, Oxford, UK. Procedures were documented in detail at <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=155580>. We leveraged the QC metrics made available by the UK Biobank and removed samples that had mismatch between genetically inferred sex and self-reported sex, samples that had high genotype missingness or extreme heterozygosity not explained by mixed ancestry or increased levels of marriage between close relatives, and one individual from each pair of the samples that were 3rd degree or more closely related relatives. We restricted our analysis to subjects that were self-reported white British and confirmed by principal component analysis (PCA) to be Caucasians. We further filtered out genetic markers that had high missing rate (>1%), low minor allele frequency (<1%), significant deviation from Hardy-Weinberg equilibrium ($p < 1e-7$), and subjects that had high missing genotype rate (>1%). 108,158 subjects (age 40–73 years; female 52.84%) and 486,175 SNPs remained for analysis after QC. [S4 Fig](#) shows the age distribution of the subjects that passed QC. The genetic similarity matrix was computed using all genotyped autosomal

SNPs. All genetic analyses were performed using PLINK 1.9 (<https://www.cog-genomics.org/plink2>) [68].

Phenotypic variables

We analyzed every trait available to us that had a sufficient sample size to produce a heritability estimate with its standard error smaller than 0.1. The traits can be classified into the following 11 domains as defined by the UK Biobank: cognitive functions, early life factors, health and medical history, life style, physical measures, psychosocial factors, sex-specific factors and sociodemographics. For continuous traits, we excluded samples that were more than 5 standard deviations away from the population mean to avoid extreme outliers and data recording errors. We only analyzed binary traits that had prevalence greater than 1% in the sample, so that we had enough statistical power to get reliable heritability estimates. We typically binarized categorical variables at a meaningful threshold close to the median and then analyzed them as binary traits. For the specific cutoff-points used to binarize each categorical variable, see [S1 Table](#).

We also analyzed a large number of self-reported illness codes and hospital in-patient diagnosis codes. Self-reported cancer and non-cancer illness codes were obtained through a verbal interview by a trained nurse at the UK Biobank assessment center on past and current medical conditions. Hospital in-patient diagnoses were obtained through medical records and were coded according to the International Classification of Diseases version-10 (ICD-10). Disease codes for each domain (self-reported cancer, self-reported non-cancer illness, and ICD-10) were organized in a hierarchical tree structure; codes closer to the root of the tree are often less specific and have larger prevalence, while codes closer to the leaves are more specific but have lower prevalence. We analyzed every disease code that had prevalence greater than 1% in the sample. 15 ICD-10 codes were excluded due to small sample sizes and large standard errors (>0.1) on heritability estimates (N70-N77 Inflammatory diseases of female pelvic organs; O20-O29 Other maternal disorders predominantly related to pregnancy; O80-O84 Delivery; N50 Other disorders of male genital organs; N80 Endometriosis; N81.1 Cystocele; N81.2 Incomplete uterovaginal prolapse; N83 Noninflammatory disorders of ovary—Fallopian tube and broad ligament; N83.2 Other and unspecified ovarian cysts; N84.1 Polyp of cervix uteri; N92.1 Excessive and frequent menstruation with irregular cycle; N93 Other abnormal uterine and vaginal bleeding; O68 Labour and delivery complicated by foetal stress [distress]; O70.1 Second degree perineal laceration during delivery; O80 Single spontaneous delivery). We also employed a data-driven approach to determine if a disease is sex-specific. More specifically, if the sample prevalence of a disease in males was more than 100 times larger than the sample prevalence in females, we defined the disease as male-specific and the analysis was restricted to males. The same approach was used to find female-specific diseases. See [S2 Table](#) for all the disease codes we analyzed.

Heritability estimation

We consider the linear random effect model $\mathbf{y} = \mathbf{g} + \mathbf{e}$, where an N -dimensional trait \mathbf{y} is partitioned into the sum of additive genetic effects \mathbf{g} and unique (subject-specific) environmental effects \mathbf{e} . The covariance structure of \mathbf{y} is $\text{cov}[\mathbf{y}] = \sigma_g^2 \mathbf{K} + \sigma_e^2 \mathbf{I}$, where \mathbf{K} is the empirical genetic similarity matrix for each pair of individuals estimated from genome-wide SNP data [4,5], \mathbf{I} is an identity matrix, σ_g^2 and σ_e^2 are the total additive genetic variance captured by genotyped common SNPs and the variance of unique environmental factors across individuals, respectively. SNP heritability is then defined as $h_{\text{SNP}}^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2) = \sigma_g^2 / \sigma_p^2$, which measures the

total phenotypic variance σ_p^2 that can be explained by total additive genetic variance tagged by genotyped SNPs, and is a lower bound for the narrow-sense heritability h^2 . When covariates need to be incorporated into the model, i.e., $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{g} + \mathbf{e}$, where \mathbf{X} is an $N \times q$ covariate matrix and $\boldsymbol{\beta}$ is a vector of fixed effects, an $N \times (N - q)$ matrix \mathbf{U} always exists, which satisfies $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{U}\mathbf{U}^T = \mathbf{P}_0$, $\mathbf{U}^T\mathbf{X} = \mathbf{0}$, and $\mathbf{P}_0 = \mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$. Applying \mathbf{U}^T to both sides of the model removes the covariate matrix [69,70].

To obtain unbiased estimates of σ_g^2 and σ_e^2 , we used a computationally efficient moment-matching approach [70,71], which is closely related to the Haseman-Elston regression [14–16] and phenotype-correlation genetic-correlation (PCGC) regression [10], and mathematically equivalent to the LD score regression under certain conditions [8,18]. Specifically, we regress the empirical estimate of the phenotypic covariance onto the matrices \mathbf{K} and \mathbf{I} : $\text{vec}[\mathbf{y}\mathbf{y}^T] = \sigma_g^2\text{vec}[\mathbf{K}] + \sigma_e^2\text{vec}[\mathbf{I}] + \epsilon$, where $\text{vec}[\cdot]$ is the matrix vectorization operator that converts a matrix into a vector by stacking its columns, and ϵ is the residual of the regression. The ordinary least squares (OLS) estimator of this multiple regression problem can be explicitly written as $\hat{\sigma}_g^2 = \frac{1}{v_K} \mathbf{y}^T(\mathbf{K} - \tau\mathbf{I})\mathbf{y}$ and $\hat{\sigma}_e^2 = \frac{1}{v_K} \mathbf{y}^T(\kappa\mathbf{I} - \tau\mathbf{K})\mathbf{y}$, where $\tau = \text{tr}[\mathbf{K}]/N$, $\kappa = \text{tr}[\mathbf{K}^2]/N$, and $v_K = N(\kappa - \tau^2)$. SNP heritability is then estimated as $\hat{h}_{\text{SNP}}^2 = \hat{\sigma}_g^2 / (\hat{\sigma}_g^2 + \hat{\sigma}_e^2) = \hat{\sigma}_g^2 / \hat{\sigma}_p^2$.

To estimate the sampling variance of \hat{h}_{SNP}^2 , we follow Visscher et al. [17] and make two assumptions: (1) the off-diagonal elements in the empirical genetic similarity matrix \mathbf{K} are small, such that $\mathbf{K} \approx \mathbf{I}$ and $\mathbf{V} = \text{cov}[\mathbf{y}] = \sigma_g^2\mathbf{K} + \sigma_e^2\mathbf{I} \approx \sigma_p^2\mathbf{I}$; and (2) the phenotypic variance σ_p^2 can be estimated with very high precision. We thus have $\text{var}[\hat{\sigma}_g^2] = \frac{2}{v_K} \text{tr}[(\mathbf{K} - \tau\mathbf{I})\mathbf{V}(\mathbf{K} - \tau\mathbf{I})\mathbf{V}] \approx 2\sigma_p^4/v_K$, and $\text{var}[\hat{h}_{\text{SNP}}^2] \approx 2/v_K$. This estimator coincides with existing results in the literature [17]. We note that the calculation of the variance of $\hat{\sigma}_g^2$ relies on an additional assumption that the trait \mathbf{y} is Gaussian distributed and thus may be suboptimal for binary traits. However, Visscher and colleagues have empirically shown that this sampling variance approximation is accurate for both continuous and binary traits when the sample size is large [17] (also see S3 Fig).

We note that for large sample size N , the $N \times N$ genetic similarity matrix \mathbf{K} and residual forming matrix \mathbf{P}_0 can be very large, making the computation memory intensive. We have developed a memory efficient algorithm that can iteratively load columns (or block columns) of \mathbf{K} into the memory to compute the SNP heritability estimate, and does not need to explicitly compute \mathbf{P}_0 or any other $N \times N$ matrices. See S1 Text for details. Matlab and Python implementations of the algorithm are available at <https://github.com/chaiyanchen/mmhe>.

For binary traits, the above calculation gives a heritability estimate on the observed scale, which is dependent on prevalence of the trait in the population. We transformed this heritability estimate to the underlying liability scale under the assumption of a classical liability threshold model [72,73], which makes heritability estimates independent of prevalence and thus comparable across traits. Specifically, heritability estimate on the liability scale can be obtained using a linear transformation of the heritability on the observed scale: $\hat{h}_{\text{SNP,L}}^2 = c\hat{h}_{\text{SNP}}^2$, where $c = P(1 - P)/\varphi(t)^2$, P is the population prevalence, $t = \Phi^{-1}(1 - P)$ is the liability threshold, Φ is the cumulative distribution function of the standard normal distribution, and φ is the density function of the standard normal distribution [3,74]. Since the UK Biobank is not designed to be ascertained for particular diseases, we assumed that population prevalence is identical to sample prevalence. The sampling variance of the heritability estimate can be transformed accordingly: $\text{var}[\hat{h}_{\text{SNP,L}}^2] = c^2\text{var}[\hat{h}_{\text{SNP}}^2]$.

Statistical analysis

In all heritability analyses, we included genotyping array, UK Biobank assessment center, age at recruitment and top 10 principal components (PCs) of the genotype data as covariates. Other covariates such as sex and handedness (e.g., when analyzing the grip strength of the left/right hand) were adjusted where appropriate. See [S1 Table](#) for the set of covariates we included in the model when estimating the heritability for each trait. To compute PCs of the genotype data, we performed pairwise linkage disequilibrium (LD) based SNP pruning at $R^2 > 0.02$ and excluded SNPs in the major histocompatibility complex (MHC) region (chr6:25-35Mb) and chromosome 8 inversion (chr8:7-13Mb). Top PCs were then computed using flashPCA [75] on the pruned data, which employs an efficient randomized algorithm and is thus scalable to large data sets with hundreds of thousands of individuals.

To examine how heritability estimates are modified by sex, we estimated heritability for each non-sex-specific trait in males and females separately. For binary traits, sample prevalence was calculated in each sex. To test if heritability estimates are significantly different by sex, we assumed that the two SNP heritability estimates to be contrasted, \hat{h}_A and \hat{h}_B , are independent and approximately Gaussian distributed, and computed the z -score of their difference: $z = (\hat{h}_A - \hat{h}_B) / \sqrt{\hat{se}_A^2 + \hat{se}_B^2}$, where \hat{se}_A^2 and \hat{se}_B^2 are standard error estimates of \hat{h}_A and \hat{h}_B , respectively. A p -value can then be computed as $p = 2 \cdot \Phi(-|z|)$, where Φ is the cumulative distribution function of the standard normal distribution.

To examine whether SNP heritability estimates vary with age, we used a sliding window approach and estimated heritability for every age range of 10 years (i.e., 40–49 years, 41–50 years, . . . , 64–73 years) by stratifying the samples. For binary traits, sample prevalence was calculated in each age range separately. We assessed whether heritability estimates exhibited a linear trend with age by fitting a regression model, $\hat{h}_k = \alpha + \overline{\text{age}}_k \gamma + \epsilon_k$, where \hat{h}_k is the heritability estimate in the k -th age range, $\overline{\text{age}}_k$ is the mean of the age range, α is an intercept, γ is the slope and ϵ_k is the residual of the regression, and testing whether γ is significantly different from zero. We weighted heritability estimates by the inverse of their standard errors when fitting the regression model, and thus put more emphasis on estimates with better precision. We only analyzed physical and cognitive measures, and did not consider disease codes and medical history in age stratification analyses because age at recruitment does not reflect disease onset.

Similarly, we used a sliding window approach to estimate the SNP heritability for each trait from the bottom 1/3 quantile to the top 1/3 quantile of the Townsend deprivation index at recruitment, a measure of material deprivation within the population of a given area. For binary traits, sample prevalence was calculated in each SES bin separately. For traits that do not reflect the status of participants at the time of recruitment (e.g., medical history and early-life factors), we have implicitly made an assumption that the SES of participants had not changed dramatically throughout their lives.

To account for multiple testing in our stratification analyses, we corrected the p -values using the effective number of independent traits we analyzed. Specifically, for each stratification analysis (sex, age and SES), we calculated the Pearson correlation coefficient for each pair of the traits using their overlapping samples. The correlation between traits that had no sample overlap, e.g., male- and female-specific factors, was set to zero. We then conducted a principal component analysis (PCA) to the constructed phenotypic correlation matrix, and estimated the effective numbers of independent traits that explained 99% of the total phenotypic variation in sex, age and SES stratification analyses to be 400, 31 and 440, respectively. Finally, we multiplied uncorrected p -values by the corresponding effective number of independent traits to obtain corrected p -values.

Supporting information

S1 Text. The moment-matching method for SNP heritability analysis.

(DOCX)

S1 Table. The SNP heritability estimates, standard error estimates, sample sizes, covariates adjusted, prevalence in the sample (for binary traits) and other relevant information for all non-disease traits analyzed in the UK Biobank. SNP Heritability estimates stratified by sex and socioeconomic status (SES) measured by the Townsend deprivation index are also presented.

(XLSX)

S2 Table. The SNP heritability estimates, standard error estimates, sample sizes, prevalence in the sample and other relevant information for all self-reported illness codes and ICD-10 codes analyzed in the UK Biobank. SNP Heritability estimates stratified by sex and socioeconomic status (SES) measured by the Townsend deprivation index are also presented.

(XLSX)

S3 Table. SNP Heritability estimates stratified by age.

(XLSX)

S4 Table. Traits whose SNP heritability estimates decrease by 0.2 or more when the major histocompatibility complex (MHC) region (chr6:25-35Mb) is removed when computing the genetic similarity matrix.

(XLSX)

S1 Fig. (A) A breakdown of the 79 self-reported non-cancer illness codes into different functional domains; (B) A breakdown of the 194 ICD-10 codes into different functional domains.

(TIFF)

S2 Fig. The mean and standard deviation (shaded region) of the 12 traits whose SNP heritability estimates significantly decrease with age.

(TIFF)

S3 Fig. A comparison of the SNP heritability and standard error (SE) estimates from the moment-matching method, genome-wide complex trait analysis (GCTA) and LD score regression (with constrained intercept and the regression weight set to the reciprocal of the LD score), by randomly sampling 10,000 subjects for each trait in [S1 Table](#).

(TIFF)

S4 Fig. Age distribution of the sample analyzed in the UK Biobank.

(TIFF)

Author Contributions

Conceptualization: TG MRS JWS.

Data curation: TG JWS.

Formal analysis: TG CYC.

Funding acquisition: JWS MRS.

Methodology: TG CYC MRS JWS.

Resources: MRS JWS.

Software: TG CYC.

Supervision: JWS MRS.

Validation: TG.

Visualization: TG.

Writing – original draft: TG JWS.

Writing – review & editing: TG CYC BMN MRS JWS.

References

1. Visscher P. M., Hill W. G., Wray N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics*, 9(4), 255–266. <https://doi.org/10.1038/nrg2322> PMID: 18319743
2. Polderman T. J., Benyamin B., De Leeuw C. A., Sullivan P. F., Van Bochoven A., Visscher P. M., et al. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47(7), 702–709. <https://doi.org/10.1038/ng.3285> PMID: 25985137
3. Lee S. H., Wray N. R., Goddard M. E., Visscher P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *The American Journal of Human Genetics*, 88(3), 294–305. <https://doi.org/10.1016/j.ajhg.2011.02.002> PMID: 21376301
4. Yang J., Benyamin B., McEvoy B. P., Gordon S., Henders A. K., Nyholt D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7), 565–569. <https://doi.org/10.1038/ng.608> PMID: 20562875
5. Yang J., Lee S. H., Goddard M. E., Visscher P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1), 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011> PMID: 21167468
6. Turkheimer E., Haley A., Waldron M., D'Onofrio B., Gottesman I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14(6), 623–628. <https://doi.org/10.1046/j.0956-7976.2003.psci.1475.x> PMID: 14629696
7. Sudlow C., Gallacher J., Allen N., Beral V., Burton P., Danesh J., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*, 12(3), e1001779. <https://doi.org/10.1371/journal.pmed.1001779> PMID: 25826379
8. Bulik-Sullivan B. K., Loh P. R., Finucane H. K., Ripke S., Yang J., Patterson N., et al. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47(3), 291–295. <https://doi.org/10.1038/ng.3211> PMID: 25642630
9. Chatterjee N., Shi J., García-Closas M. (2016). Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nature Reviews Genetics*, 17(7), 392–406. <https://doi.org/10.1038/nrg.2016.27> PMID: 27140283
10. Golan D., Lander E. S., Rosset S. (2014). Measuring missing heritability: inferring the contribution of common variants. *Proceedings of the National Academy of Sciences*, 111(49), E5272–E5281.
11. Finucane H. K., Bulik-Sullivan B., Gusev A., Trynka G., Reshef Y., Loh P. R., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, 47(11), 1228–1235. <https://doi.org/10.1038/ng.3404> PMID: 26414678
12. Bulik-Sullivan B., Finucane H. K., Anttila V., Gusev A., Day F. R., Loh P. R., et al. (2015a). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, 47(11), 1236–1241.
13. Zheng J., Erzurumluoglu A. M., Elsworth B. L., Kemp J. P., Howe L., Haycock P. C., et al. (2016). LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*, btw613.
14. Elston R. C., Buxbaum S., Jacobs K. B., Olson J. M. (2000). Haseman and Elston revisited. *Genetic Epidemiology*, 19(1), 1–17. [https://doi.org/10.1002/1098-2272\(200007\)19:1<1::AID-GEPI1>3.0.CO;2-E](https://doi.org/10.1002/1098-2272(200007)19:1<1::AID-GEPI1>3.0.CO;2-E) PMID: 10861893
15. Haseman J. K., Elston R. C. (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behavior Genetics*, 2(1), 3–19. PMID: 4157472
16. Sham P. C., Purcell S. (2001). Equivalence between Haseman-Elston and variance-components linkage analyses for sib pairs. *The American Journal of Human Genetics*, 68(6), 1527–1532. <https://doi.org/10.1086/320593> PMID: 11353401

17. Visscher P. M., Hemani G., Vinkhuyzen A. A., Chen G. B., Lee S. H., Wray N. R., et al. (2014). Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *PLoS Genetics*, 10(4), e1004269. <https://doi.org/10.1371/journal.pgen.1004269> PMID: 24721987
18. Bulik-Sullivan, B. (2015). Relationship between LD score and Haseman-Elston regression. *bioRxiv*, 018283.
19. MacGregor S., Cornes B. K., Martin N. G., Visscher P. M. (2006). Bias, precision and heritability of self-reported and clinically measured height in Australian twins. *Human Genetics*, 120(4), 571–580. <https://doi.org/10.1007/s00439-006-0240-z> PMID: 16933140
20. Silventoinen K., Magnusson P. K., Tynelius P., Kaprio J., Rasmussen F. (2008). Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. *Genetic Epidemiology*, 32(4), 341–349. <https://doi.org/10.1002/gepi.20308> PMID: 18271028
21. Silventoinen K., Sammalisto S., Perola M., Boomsma D. I., Cornes B. K., Davis C., et al. (2003). Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Research*, 6(05), 399–408.
22. Visscher P. M., McEvoy B., Yang J. (2010). From Galton to GWAS: quantitative genetics of human height. *Genetics Research*, 92, 371–379. <https://doi.org/10.1017/S0016672310000571> PMID: 21429269
23. Zaitlen N., Kraft P., Patterson N., Pasaniuc B., Bhatia G., Pollack S., et al. (2013). Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genetics*, 9(5), e1003520. <https://doi.org/10.1371/journal.pgen.1003520> PMID: 23737753
24. Zaitlen N., Pasaniuc B., Sankararaman S., Bhatia G., Zhang J., Gusev A., et al. (2014). Leveraging population admixture to characterize the heritability of complex traits. *Nature Genetics*, 46(12), 1356–1362. <https://doi.org/10.1038/ng.3139> PMID: 25383972
25. Zuk O., Hechter E., Sunyaev S. R., Lander E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, 109(4), 1193–1198.
26. Vattikuti S., Guo J., Chow C. C. (2012). Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. *PLoS Genetics*, 8(3), e1002637. <https://doi.org/10.1371/journal.pgen.1002637> PMID: 22479213
27. Yang J., Bakshi A., Zhu Z., Hemani G., Vinkhuyzen A. A., Lee S. H., et al. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature Genetics*, 47(10), 1114–1120. <https://doi.org/10.1038/ng.3390> PMID: 26323059
28. Bates T. C., Lewis G. J., Weiss A. (2013). Childhood socioeconomic status amplifies genetic effects on adult intelligence. *Psychological Science*, 24(10), 2111–2116. <https://doi.org/10.1177/0956797613488394> PMID: 24002887
29. Hanscombe K. B., Trzaskowski M., Haworth C. M., Davis O. S., Dale P. S., Plomin R. (2012). Socioeconomic status (SES) and children's intelligence (IQ): In a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS One*, 7(2), e30320. <https://doi.org/10.1371/journal.pone.0030320> PMID: 22312423
30. Kirkpatrick R. M., McGue M., Iacono W. G. (2015). Replication of a gene–environment interaction via multimodel inference: additive-genetic variance in Adolescents' General Cognitive Ability Increases with Family-of-Origin Socioeconomic Status. *Behavior Genetics*, 45(2), 200–214. <https://doi.org/10.1007/s10519-014-9698-y> PMID: 25539975
31. Turkheimer E., Beam C. E., Davis D. W. (2015). The Scarr-Rowe interaction in complete seven-year WISC data from the Louisville twin study: Preliminary report. *Behavior Genetics*, 45(6), 635–639. <https://doi.org/10.1007/s10519-015-9760-4> PMID: 26497158
32. Haworth C. M., Wright M. J., Luciano M., Martin N. G., De Geus E. J. C., Van Beijsterveldt C. E. M., et al. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular psychiatry*, 15(11), 1112–1120. <https://doi.org/10.1038/mp.2009.55> PMID: 19488046
33. Davis O. S., Haworth C. M., Plomin R. (2009). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science*, 20(10), 1301–1308. <https://doi.org/10.1111/j.1467-9280.2009.02433.x> PMID: 19732386
34. Tucker-Drob E. M., Bates T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological science*, 0956797615612727.
35. Rietveld C. A., Esko T., Davies G., Pers T. H., Turley P., Benyamin B., et al. (2014). Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proceedings of the National Academy of Sciences*, 111(38), 13790–13794.

36. Zillikens M. C., Yazdanpanah M., Pardo L. M., Rivadeneira F., Aulchenko Y. S., Oostra B. A., et al. (2008). Sex-specific genetic effects influence variation in body composition. *Diabetologia*, 51(12), 2233–2241. <https://doi.org/10.1007/s00125-008-1163-0> PMID: 18839131
37. Hottenga J. J., Boomsma D. I., Kupper N., Posthuma D., Snieder H., Willemsen G., et al. (2005). Heritability and stability of resting blood pressure. *Twin Research and Human Genetics*, 8(05), 499–508.
38. MacGregor A. J., Snieder H., Rigby A. S., Koskenvuo M., Kaprio J., Aho K., et al. (2000). Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis & Rheumatism*, 43(1), 30.
39. Brown W. M., Beck S. R., Lange E. M., Davis C. C., Kay C. M., Langefeld C. D., et al. (2003). Age-stratified heritability estimation in the Framingham Heart Study families. *BMC Genetics*, 4(1), 1.
40. Ortega-Alonso A., Sipilä S., Kujala U. M., Kaprio J., Rantanen T. (2009). Genetic influences on change in BMI from middle to old age: a 29-year follow-up study of twin sisters. *Behavior Genetics*, 39(2), 154–164. <https://doi.org/10.1007/s10519-008-9245-9> PMID: 19067155
41. Schousboe K., Visscher P. M., Erbas B., Kyvik K. O., Hopper J. L., Henriksen J. E., et al. (2004). Twin study of genetic and environmental influences on adult body size, shape, and composition. *International Journal of Obesity*, 28(1), 39–48. <https://doi.org/10.1038/sj.jco.0802524> PMID: 14610529
42. Haworth C., Carnell S., Meaburn E. L., Davis O. S., Plomin R., Wardle J. (2008). Increasing heritability of BMI and stronger associations with the FTO gene over childhood. *Obesity*, 16(12), 2663–2668. <https://doi.org/10.1038/oby.2008.434> PMID: 18846049
43. Min J., Chiu D. T., Wang Y. (2013). Variation in the heritability of body mass index based on diverse twin studies: a systematic review. *Obesity Reviews*, 14(11), 871–882. <https://doi.org/10.1111/obr.12065> PMID: 23980914
44. Coultas D. B., Hanis C. L., Howard C. A., Skipper B. J., Samet J. M. (1991). Heritability of ventilatory function in smoking and nonsmoking New Mexico Hispanics. *American Review of Respiratory Disease*, 144(4), 770–775. <https://doi.org/10.1164/ajrccm/144.4.770> PMID: 1928947
45. Hottenga J. J., Whitfield J. B., De Geus E. J., Boomsma D. I., Martin N. G. (2006). Heritability and stability of resting blood pressure in Australian twins. *Twin Research and Human Genetics*, 9(02), 205–209.
46. McClearn G. E., Svartengren M., Pedersen N. L., Heller D. A., Plomin R. (1994). Genetic and environmental influences on pulmonary function in aging Swedish twins. *Journal of Gerontology*, 49(6), M264–M268.
47. Menni C., Mangino M., Zhang F., Clement G., Snieder H., Padmanabhan S., et al. (2013). Heritability analyses show visit-to-visit blood pressure variability reflects different pathological phenotypes in younger and older adults: evidence from UK twins. *Journal of Hypertension*, 31(12), 2356–2361. <https://doi.org/10.1097/HJH.0b013e32836523c1> PMID: 24029873
48. Tropf F. C., Barban N., Mills M. C., Snieder H., Mandemakers J. J. (2015). Genetic influence on age at first birth of female twins born in the UK, 1919–68. *Population studies*, 69(2), 129–145. <https://doi.org/10.1080/00324728.2015.1056823> PMID: 26234944
49. Vinck W. J., Fagard R. H., Loos R., Vlietinck R. (2001). The impact of genetic and environmental influences on blood pressure variance across age-groups. *Journal of Hypertension*, 19(6), 1007–1013. PMID: 11403347
50. Wang B., Liao C., Zhou B., Cao W., Lv J., Yu C., et al. (2015). Genetic contribution to the variance of blood pressure and heart rate: A systematic review and meta-regression of twin studies. *Twin Research and Human Genetics*, 18(02), 158–170.
51. Domingue, B. W., Belsky, D. W., Harrati, A., Conley, D., Weir, D., Boardman, J. (2016). Mortality selection in a genetic sample and implications for association studies. *bioRxiv*, 049635
52. Tyrrell J., Jones S. E., Beaumont R., Astley C. M., Lovell R., Yaghootkar H., et al. (2016). Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *bmj*, 352, i582. <https://doi.org/10.1136/bmj.i582> PMID: 26956984
53. Lyall D. M., Cullen B., Allerhand M., Smith D. J., Mackay D., Evans J., et al. (2016). Cognitive test scores in UK Biobank: data reduction in 480,416 participants and longitudinal stability in 20,346 participants. *PLoS one*, 11(4), e0154222. <https://doi.org/10.1371/journal.pone.0154222> PMID: 27110937
54. Siebert S., Lyall D. M., Mackay D. F., Porter D., McInnes I. B., Sattar N., et al. (2016). Characteristics of rheumatoid arthritis and its association with major comorbid conditions: cross-sectional study of 502 649 UK Biobank participants. *RMD open*, 2(1), e000267. <https://doi.org/10.1136/rmdopen-2016-000267> PMID: 27403335
55. Castro V. M., Minnier J., Murphy S. N., Kohane I., Churchill S. E., Gainer V., et al. (2015). Validation of electronic health record phenotyping of bipolar disorder cases and controls. *American Journal of Psychiatry*, 172(4), 363–372. <https://doi.org/10.1176/appi.ajp.2014.14030423> PMID: 25827034

56. Ford E., Carroll J. A., Smith H. E., Scott D., Cassell J. A. (2016). Extracting information from the text of electronic medical records to improve case detection: a systematic review. *Journal of the American Medical Informatics Association*, ocv180.
57. Liao K. P., Cai T., Savova G. K., Murphy S. N., Karlson E. W., Ananthakrishnan A. N., et al. (2015). Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ*, 350, h1885. <https://doi.org/10.1136/bmj.h1885> PMID: 25911572
58. Perlis R. H., Iosifescu D. V., Castro V. M., Murphy S. N., Gainer V. S., Minnier J., et al. (2012). Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. *Psychological Medicine*, 42(1), 41–50. <https://doi.org/10.1017/S0033291711000997> PMID: 21682950
59. Speed D., Hemani G., Johnson M. R., Balding D. J. (2012). Improved heritability estimation from genome-wide SNPs. *The American Journal of Human Genetics*, 91(6), 1011–1021. <https://doi.org/10.1016/j.ajhg.2012.10.010> PMID: 23217325
60. Lee S. H., Yang J., Chen G. B., Ripke S., Stahl E. A., Hultman C. M., et al. (2013). Estimation of SNP heritability from dense genotype data. *The American Journal of Human Genetics*, 93(6), 1151–1155. <https://doi.org/10.1016/j.ajhg.2013.10.015> PMID: 24314550
61. Yang J., Manolio T. A., Pasquale L. R., Boerwinkle E., Caporaso N., Cunningham J. M., et al. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature genetics*, 43(6), 519–525. <https://doi.org/10.1038/ng.823> PMID: 21552263
62. Tropf, F. C., Verweij, R. M., van der Most, P. J., Stulp, G., Bakshi, A., Briley, D. A., et al. (2016). Mega-analysis of 31,396 individuals from 6 countries uncovers strong gene-environment interaction for human fertility. *BioRxiv*, 049163.
63. Yang J., Bakshi A., Zhu Z., Hemani G., Vinkhuyzen A. A., Nolte I. M., et al. (2015). Genome-wide genetic homogeneity between sexes and populations for human height and body mass index. *Human molecular genetics*, ddv443.
64. Speed D., Balding D. J. (2015). Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics*, 16(1), 33–44. <https://doi.org/10.1038/nrg3821> PMID: 25404112
65. Kumar S. K., Feldman M. W., Rehkopf D. H., Tuljapurkar S. (2016). Limitations of GCTA as a solution to the missing heritability problem. *Proceedings of the National Academy of Sciences*, 113(1), E61–E70.
66. Yang J., Lee S. H., Wray N. R., Goddard M. E., Visscher P. M. (2016). GCTA-GREML accounts for linkage disequilibrium when estimating genetic variance from genome-wide SNPs. *Proceedings of the National Academy of Sciences*, 113(32), E4579–E4580.
67. Kumar S. K., Feldman M. W., Rehkopf D. H., Tuljapurkar S. (2016). Reply to Yang et al.: GCTA produces unreliable heritability estimates. *Proceedings of the National Academy of Sciences*, 201608425.
68. Chang C. C., Chow C. C., Tellier L. C., Vattikuti S., Purcell S. M., Lee J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4(1), 1.
69. Ge T., Nichols T. E., Lee P. H., Holmes A. J., Roffman J. L., Buckner R. L., et al. (2015). Massively expedited genome-wide heritability analysis (MEGHA). *Proceedings of the National Academy of Sciences*, 112(8), 2479–2484.
70. Ge T., Reuter M., Winkler A. M., Holmes A. J., Lee P. H., Tirrell L. S., et al. (2016). Multidimensional heritability analysis of neuroanatomical shape. *Nature Communications*, 7, 13291. <https://doi.org/10.1038/ncomms13291> PMID: 27845344
71. Zhou, X. (2016). A unified framework for variance component estimation with summary statistics in genome-wide association studies. *bioRxiv*, 042846.
72. Falconer D. S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics*, 29(1), 51–76.
73. Pearson K., Lee A. (1901). Mathematical contributions to the theory of evolution VII—On the application of certain formulae in the theory of correlation to the inheritance of characters not capable of quantitative measurement. *Proceedings of the Royal Society of London*, 66(424–433), 324–327.
74. Dempster E. R., Lerner I. M. (1950). Heritability of threshold characters. *Genetics*, 35(2), 212. PMID: 17247344
75. Abraham G., Inouye M. (2014). Fast principal component analysis of large-scale genome-wide data. *PLoS One*, 9(4), e93766. <https://doi.org/10.1371/journal.pone.0093766> PMID: 24718290