medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Genetic analyses on the health impacts of testosterone highlights effects on female-specific diseases and sex differences

Jaakko T. Leinonen[1], Nina Mars[1]*, Leevi E. Lehtonen[1]*, Ari Ahola-Olli[1] *, Sanni Ruotsalainen[1], Terho Lehtimäki[2,3], Mika Kähönen[3,4], Olli Raitakari[5,6,7], Mark Daly[1,8], Tiinamaija Tuomi[1,9,10,11,12], Samuli Ripatti[1,8,13], Matti Pirinen[1,13,14], Taru Tukiainen[1]

1 = Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; 2 = Department of Clinical Chemistry, Finnish Cardiovascular Research Center – Tampere; 3 = Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland; 4 = Department of Clinical Physiology, Finnish Cardiovascular Research Center – Tampere; 5 = Centre for Population Health Research, University of Turku and Turku University Hospital; 6 = Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku; 7 = Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; 8 = Broad Institute of MIT and Harvard, Cambridge, MA, USA, 9 = Department of Endocrinology, Abdominal Centre, Helsinki University Hospital, Helsinki, Finland, 10 = Folkhalsan Research Center, Helsinki, Finland, 11 = Research Programs Unit, Clinical and Molecular Metabolism, University of Helsinki, Helsinki, Finland, 12 = Lund University Diabetes Centre, Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden, 13 = Department of Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland, 14 = Helsinki Institute for Information Technology HIIT and Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland. * Authors contributed equally to this work. Correspondence: taru.tukiainen@helsinki.fi

Abstract

Testosterone (T) is linked with diverse characteristics of human health, yet, whether these associations reflect correlation or causation remains debated. Here, we provide a broad perspective on the role of T on complex diseases in both sexes leveraging genetic and health registry data from the UK Biobank and FinnGen (total N=625,650).

We find genetically predicted T affects sex-biased and sex-specific traits, with a particularly pronounced impact on female reproductive health. We show T levels are intricately involved in metabolism, sharing many associations with sex hormone binding globulin (SHBG), but report lack of direct causality behind most of these associations. Across other disease domains, including behavior, we find little evidence for a significant contribution from normal variation in T levels. Highlighting T's unique biology, we show T displays antagonistic effects on stroke risk and reproduction in males and females.

Overall, we underscore the involvement of T in both male and female health, and the complex mechanisms linking T levels to disease risk and sex differences.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Introduction

Testosterone (T) is the male sex hormone responsible for regulation of development of primary and secondary male sexual characteristics. Individual variation in T levels has been suggested to shape human physiology broadly, including effects on disease risk in both males and females (1-3). Epidemiological studies and randomized clinical trials for T replacement therapy have observed associations between serum T levels and various traits ranging from type 2 diabetes (T2D) and cardiovascular disease to body composition and behavior (1-11). Yet, these studies have yielded partly mixed results, and, in many instances, the proposed relationships between T, complex traits and disease remain elusive (2, 7-11).

Besides the disease links, T is a known driver for sex differences. After puberty, males and females differ extensively with respect to their average T levels, with males showing roughly 7-15-fold higher serum total T concentrations (1, 12). This difference largely results from the testicular T production in males that far exceeds the amount of T produced in the ovaries and the adrenal gland in females, and is known to directly contribute to variation in, for instance, body composition between the sexes (1, 12).

In the human body, the majority of T is bound to a carrier molecule, whereas only a small fraction (1-3%) of this total T exists as free T, considered to represent the most potent form of T in terms of biological activity (13-15). Most of the remaining T in circulation is tightly bound by sex hormone-binding globulin (SHBG), and the bulk of the rest remains attached to carrier proteins like serum albumin (13, 14). Non-SHBG bound T is often approximated with free androgen index (FAI) (Figure 1A) (4, 13, 14).

In addition to processes affecting T production, circulating T levels are determined in both sexes by factors affecting T uptake in tissues, metabolism of T and the excretion of T, subject to homeostatic and allostatic regulation, allowing T levels to fluctuate based on internal and external stimuli on a daily basis (16, 17). However, twin studies indicate that heritability of serum T is relatively high in both sexes, up to 65% in males (18, 19). Until recently, few genetic variants were shown to associate with circulating T levels and related traits in the general population (4, 18, 20, 21). Yet, through the emergence of genetic and biomarker data

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

from large biobanks, currently more than a hundred loci for serum total T, SHBG and bioavailable T (i.e. free T and FAI) have been identified, with evidence for sex-specific genetic effects (22-24).

Recent efforts have used genetics to address the potential causal contribution of adult T to selected complex traits in both sexes, including T2D, body composition and hormonal cancers, or have examined the effects of free T more broadly in males based on the UK Biobank data (22, 23, 25-27). Here, we extend these investigations combining data from the UK Biobank (N=408,186) and FinnGen (N=217,464). Leveraging the extensive healthcare registry data of FinnGen, sex-specific polygenic scores (PGS) for T traits and SHBG, and public data from genome-wide association studies (GWAS), we studied the role of T in both sexes across traits ranging from metabolic conditions and sex-specific reproductive disorders to neurological and behavioral endpoints.

Combining multiple analysis strategies, we illustrate causal links between T and sex-specific diseases and sex-biased traits, including many reproductive conditions in females, while finding that the association between T and many metabolic phenotypes may instead largely stem from shared etiology. Overall, our data provides cues into the biology of T and its contribution to disease risk and complex traits, stressing the unique genetic architecture of serum T levels, and emphasizing the crucial role of the hormone also for women's health.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Results

We utilized the rich biochemical and health information available in two population-scale genetic datasets and analysis methods building on GWAS discovery (Figure 1). In brief, we conducted sex-stratified GWAS for T, SHBG, FAI and free T, using data available in the UK Biobank (Figure 1A), from which we built sex-specific PGS for these four traits. The PGSs capture the combined genetic effects on T and SHBG levels, and therefore serve as a proxy for cumulative post-pubertal T exposure. Using an external dataset (Young Finns Study; YFS), we validated the performance of the PGSs. We then investigated the effects of the PGSs on a wide range of diseases across diverse clinical entities using the FinnGen study (Figure 1B). Lastly, we evaluated causal relationships and genetic correlations between the studied T traits and complex traits, leveraging publicly available GWAS summary statistics.

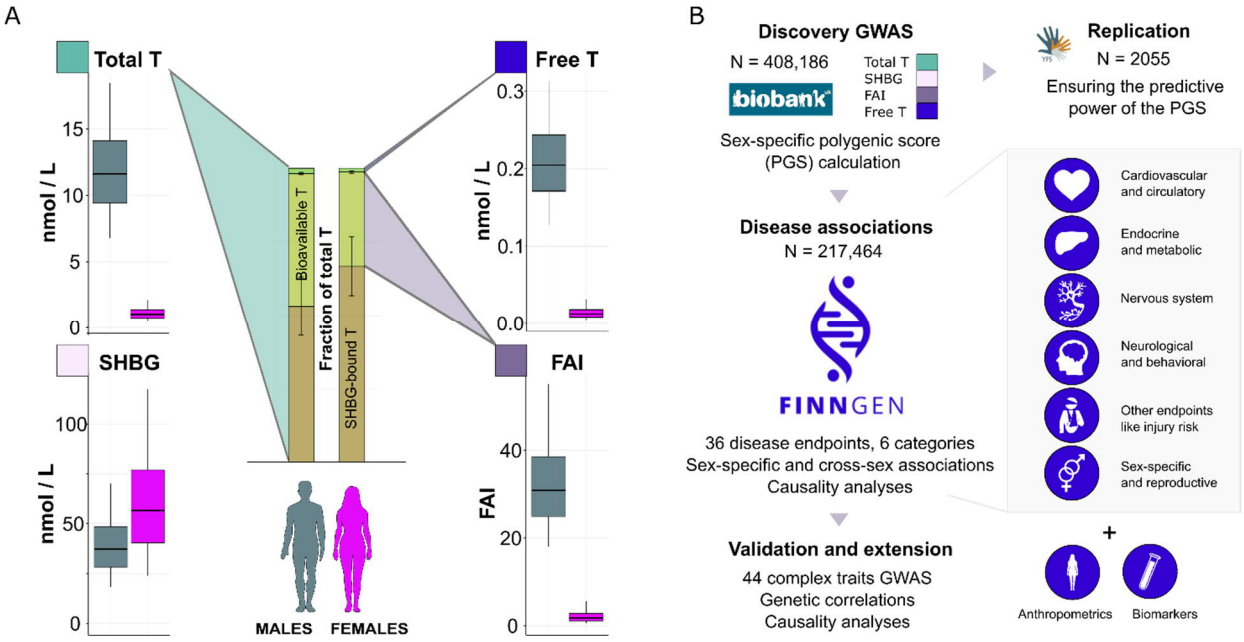


Figure 1. Illustration of the studied traits and study overview. A) Sex-specific distributions for serum total T and SHBG and calculated FAI and free T levels for the UK Biobank participants included in the genome-wide association study (GWAS). The box plots show median (black line), lower and upper quartiles (colored area of the box) and the error bars indicate 5% and 95% quantiles. B) Overview of the study design to assess the contribution of T to health and disease using genetic approaches and biobank data. We conducted the discovery GWAS in the UK Biobank, built sex-specific PGSs for the four T-related traits, validated the PGSs in the Young Finns Study (YFS), and performed complex disease and trait associations in FinnGen (release 5) and using publicly available GWAS data.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

GWAS and polygenic scores for testosterone traits

We identified more than a hundred genome-wide significant (*p*<5e-08) loci for all testosterone traits (up to 263 loci for SHBG in males) in the UK Biobank GWAS (Methods; Supplementary Tables 1-9). In both sexes, all traits displayed significant contribution from common variants (allele frequency >1%) to the trait variability, with the SNP heritability (h^2) estimates ranging from 10% for total T in females to 28% for SHBG in males (Supplementary Table 9). The associated loci were enriched for genes affecting steroid hormone biosynthesis, metabolism and excretion, with preferential expression in the liver for all the studied traits, in line with recent findings (Supplementary Figures 1&2, Supplementary Table 10) (22-24).

The loci affecting SHBG were largely shared between the sexes (genetic correlation (r_g) =0.88, *p*=9.7e-197), for FAI sharing was intermediate (r_g=0.54, *p*=5.8e-26), and we observed a near-zero genetic correlation estimates for both serum T and free T between males and females (r_g=0.08 and 0.05, respectively, *p*>0.05), indicating sex-specific genetic determinants, as previously reported (22, 23). Co-localization analyses between the male and female GWAS further confirmed the widespread sex-specificity of the genetic loci (Methods, Supplementary Figure 3 and Supplementary Tables 11-19) (23).

Reflecting the sex-specific genetic architecture, we observed strong genetic correlation between total T and SHBG only in males (r_g=0.78 in males vs. 0.05 in females, Supplementary Figure 3 and Supplementary Table 9). In females, instead, the genetic determinants for FAI and free T were shared with SHBG (e.g. r_g=-0.80 with FAI). Further highlighting the close connection between T and SHBG, we detected evidence for SHBG being causal for total T levels in males (genetic causality proportion (GCP)=0.80, *p*=5.8e-05, Methods), whereas in females SHBG appeared to control especially FAI and free T fractions (GCP=0.83, *p*=3.8e-07 for free T, Supplementary Table 9).

To study the impacts of T in datasets where T measurements are not directly available, we next constructed sex-specific genetic predictors for T levels, PGS, for each trait applying the LDpred algorithm (28) to the sex-specific GWASs (Methods). We tested the predictive ability of the PGS in the YFS where the phenotypic variance explained by the PGS (R^2) ranged between 2.5% (male free T) to 9.1% (male SHBG), indicating the

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

PGS predict T and SHBG levels in an independent cohort (Supplementary Table 20). Notably, the sex-specific PGSs for T and free T had no predictive value in the opposite sex (Supplementary Table 20 and Supplementary Figure 4).

Studying the links between cumulative T exposure and disease

We continued by using the PGS to study how post-pubertal T exposure associates with disease risk, the associations potentially implying causal relationships (28). To this end we used the FinnGen data, consisting of 217,464 (94,478 males, 122,986 females) Finnish participants with genotypes linked to up to 46 years of follow-up within nationwide healthcare registries. At present, the study includes roughly 5% of the Finnish adult population (29). We studied 36 diseases with potential links to hormones from the following categories: 1) endocrine and metabolic problems, 2) sex-specific endpoints (many specific to females, e.g. postmenopausal bleeding, (PMB)), 3) cardiovascular and circulatory system, 4) nervous system disease, 5) behavioral and neurological diagnoses and 6) other endpoints like injury risk (Supplementary Table 21). The number of cases ranged from 229 individuals diagnosed with hirsutism to 68,774 statin users.

We first tested whether the PGS are associated with disease risk in a sex-specific manner, observing 32 associations (p<0.0014, after Bonferroni correction for 36 independent tests; Supplementary Table 21 and Supplementary Figure 5). The vast majority of these associations involved endocrine, metabolic and sex-specific disorders, highlighting in particular female-specific endpoints (Figure 2A&B, and Supplementary Table 21). Further, multiple PGSs often associated with the same endpoint. In males, both total T and SHBG PGSs often associated with reduced disease risk, whereas we saw few associations to free T. In females, higher bioavailable T (FAI and free T) PGSs were linked with an increased risk for multiple diseases, often showing inverse associations to SHBG. Given these shared associations with SHBG, we additionally included the SHBG PGS as a covariate in the analyses to distinguish true T-driven effects.

Underscoring T's involvement in metabolism, larger PGS values for T and SHBG were consistently associated with reduced T2D risk and statin use in both sexes (for total T in males, HR=0.94, p=1.4e-17 and 0.96, p=1.1e-15; for total T in females, HR=0.94, p=9.4e-12 and 0.97, p=5.9e-06 per 1SD increase in PGS, T2D and statin

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

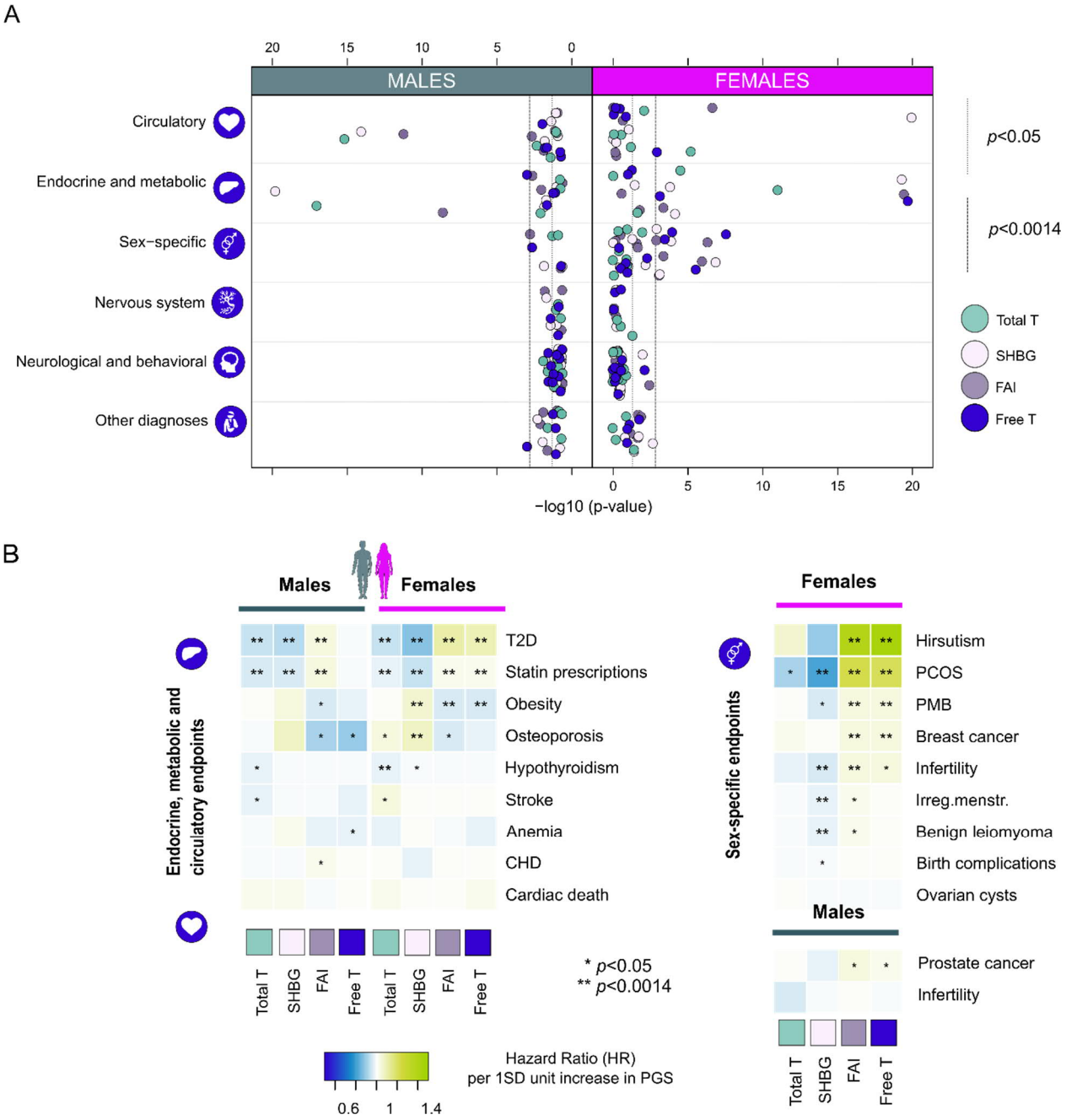
use, respectively). However, the SHBG adjusted analyses suggested these associations were not primarily attributable to androgen action, as the effect of the T was substantially attenuated in all cases (e.g. HR=1.00, p=0.93 for T and T2D in females, Supplementary Table 21 and Supplementary Figure 6). In females, higher T PGS was additionally associated with lower hypothyroidism risk (HR=0.97, *p*=3.0e-05), an association that remained after SHBG adjustment (HR=0.97, p=3.0e-04). We also detected suggestive links to bone strength and injury risk in both sexes, but with the exception of SHBG associating with osteoporosis in females (HR=1.08, *p*=0.00015), none of these findings reached statistical significance after Bonferroni correction.

In the sex-specific category, we replicated the known associations of free T to PCOS and breast cancer risk in females (22) (HR=1.02, *p*=2.8e-06 and HR=1.04, *p*=0.0001 for free T PGS) (Figure 2B and Supplementary Table 20). Of the novel endpoints, we robustly linked free T with hirsutism and post-menopausal bleeding (HR=1.45, *p*=2.7e-08 and HR=1.05, *p*=0.00032). The association to hirsutism (excessive hair growth in a male-type fashion) appeared particularly pronounced, with the risk of this condition almost doubling with a 2SD change in free T PGS. With the exception of PCOS, all these associations strengthened upon SHBG adjustment (Supplementary Table 21 and Supplementary Figure 6). Higher FAI and free T PGSs associated also with increased infertility risk in females (HR=1.04, *p*=0.00076 and HR=1.04, *p*=0.0050, respectively), but here we instead observed confounding by SHBG. In contrast to FAI and free T, higher PGS for SHBG associated with positive effects on several of the female reproductive health endpoints (e.g. HR=0.98, *p*=0.00116 for irregular menstruation). In males, higher PGSs for male free T and FAI were linked with increased prostate cancer risk in a nominally significant fashion (HR=1.03, *p*=0.0083 for free T, HR=1.03, *p*=0.0058 for FAI).

We observed no statistically significant associations to other diseases (Figure 2A). We for example detected no associations to any of the 13 neurological/behavioral endpoints studied, including Alzheimer's disease, alcohol use, and conduct and anxiety disorders (all *p*>0.0014) in both sexes. However, the PGSs did show several nominal associations to cardiovascular disease (HR=1.02, *p*=0.0077 for FAI and coronary heart disease (CHD) risk in males; HR=0.97, *p*=0.017 for T and stroke risk in males; HR=1.04, *p*=0.0080 for T and stroke risk in females; Figure 2, and Supplementary Table 21).

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

169



170

171 Figure 2. Results from the PGS associations with disease endpoints in the FinnGen. Panel A illustrates the distribution of association
172 *p*-values by disease category for each biomarker PGS separately for both sexes. For visual clarity *p*-values capped at 1e-20. Panel B
173 shows hazard ratios per one SD increase in PGS for 20 traits from endocrine, metabolic, circulatory and sex-specific categories.
174 *$p<0.05$, **$p<0.0014$, corresponding to Bonferroni correction for 36 independent traits. PCOS = polycystic ovary syndrome, PMB =
175 post-menopausal bleeding, CHD = coronary heart disease.

176

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Evaluation of causal relationships between T and disease

Although the PGS associations may imply direct causality of the first trait to another, these are also prone for confounders including genetic pleiotropy (30). We estimated causal relationships between T and the studied endpoints using two complementary MR methods that correct for potential pleiotropy: LCV (31), representing a genome-wide approach, and MR-Egger (32) that uses significantly associated SNPs. For the sex-specific PGS data, in 23/252 instances one or both methods suggested evidence for a causal relationship (p<0.0014) between a PGS and a disease (Figure 3, Supplementary Figure 7, and Supplementary Table 22). Finally, we distinguished between the effects of T and SHBG by including the latter as a covariate in multivariable MR Egger analyses (33).

The causality analyses supported the role of T in the regulation of female reproductive health. Under both MR models, we observed causality between total and free T and postmenopausal bleeding (β=0.61, p=8.6e-05 and β=0.51, p=0.00034). We also observed nominally significant evidence also for expected causal relationships between female free T, hirsutism (β=1.68, p=0.0095) and PCOS (GCP=0.54, p=0.0017) (Figure 3 and Supplementary Table 22). Additionally, both approaches indicated causality between T and hormonal cancers (e.g. GCP=-0.55, p=0.041 and β=0.34, p=0.0031 for female free T and breast cancer; GCP=-0.37, p=2.8e-16; β=0.32, p=0.014 for male free T and prostate cancer). Notably, genetically predicted free T was linked with increased cancer risk also in the opposite sex (GCP=-0.37, p=6.6e-19 for male free T and breast cancer, and GCP=-0.74, p=0.0041 for female free T and prostate cancer). Although SHBG appeared causal to irregular menstruation (GCP=0.83, p=1.8e-06, β=-0.11, p=0.022), adjusting for the effects of SHBG in the multivariable MR Egger analyses seemed to rather strengthen most causality estimates for T and free T (Supplementary Table 22).

Despite the many associations to metabolism-related endpoints (Figure 2), here the evidence for causality was limited to nominally significant links between female FAI and obesity (GCP=-0.63, p=0.014) and male T and SHBG with statin use (β=-0.22, p=0.018 and β=-0.14, p=0.016). The support for T levels being causal for

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

metabolic traits was thus sparse, and for example the link between total T and statins appeared confounded by SHBG (the effect being close to zero in the adjusted model, Supplementary Table 22). Notably, we also consistently saw no causality between T and T2D under LCV and both univariate and multivariate MR-Egger models (p>0.05 for all), a result that remained with both sex-combined and sex-specific T2D GWAS from FinnGen (Figure 3, Supplementary Figure 7 and Supplementary Table 22).

Finally, evidence for a causal relationship between T levels and disease risk was detected in some instances where the PGS were not associated with a given endpoint. Both T and SHBG were for example linked with injury risk in both sexes (e.g. GCP=0.53, p=0.0018 and β=0.17, p=1.3e-05 for SHBG increasing forearm/elbow injuries in males), potentially aligned with known hormonal contribution to bone strength (34). Notably, nominal causality between male total T and osteoporosis risk vanished when adjusting with SHBG (β=0.28, p=0.019 to β=0.034, p=0.85 Figure 3 and Supplementary Table 22). The multivariable MR Egger analyses also highlighted the potential protective effect of T on seropositive rheumatoid arthritis (RA) in females (β=-0.43, p=0.0079 for 1SD increase in total T, β=-0.61, p=4.7e-05 for FAI and β=-0.48, p=0.0025 for free T). Also, while the PGSs did not associate with any of the neurological endpoints, LCV suggested a causal relationship between higher SHBG and ADHD in both sexes, and higher free T and increased risk for conduct disorder but decreased risk for emotional instability in males (p<0.0014, Supplementary Table 22), pointing to potential hormonal involvement in the regulation of neuronal processes.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

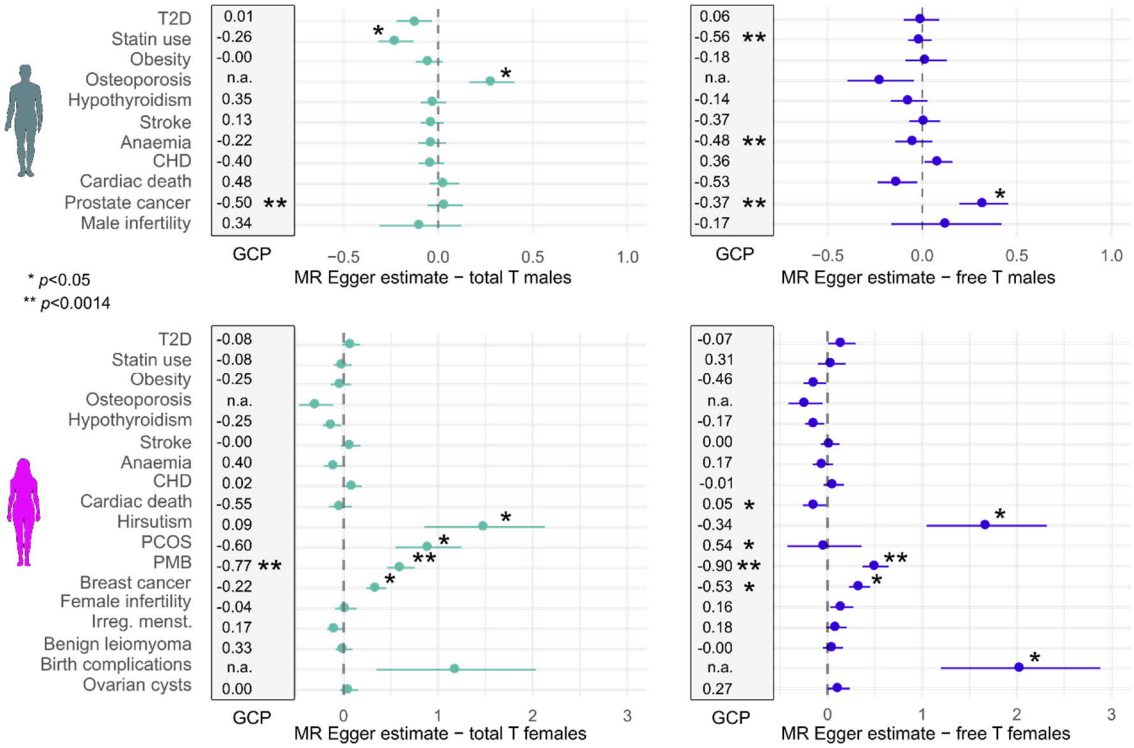


Figure 3. Results from the causality analyses in FinnGen, showing MR estimates for total and free T in males and females. The figure includes traits from the endocrine, metabolic, circulatory and sex-specific categories shown in Figure 2. Shown are both genetic causality proportion (GCP) estimates from LCV (within the grey boxes, 0=no causality, 1= fully causal, minus sign indicates suggested reverse causality i.e. the endpoint affecting T), and MR Egger betas and standard error for each trait (horizontal lines). The MR Egger value in the x-axis corresponds to SD increase in risk/per SD increase in T/free T. *p<0.05, **p<0.0014, n.a.=not applicable due to low heritability estimate in FinnGen under the LCV model.

Results from the cross-sex PGS associations in FinnGen

Owing to the unique genetic architecture, the sex-specific PGS for T and free T do not predict the corresponding hormone levels in the opposite sex (Supplementary Table 20). Given this, we reasoned that cross-sex analyses, i.e., analysis of the effect of a sex-specific PGS in the opposite sex, would provide us with an additional means to assess if the original associations stem from T action, and to detect potential antagonistic effects for the PGSs between the sexes.

We first concentrated on replicating the associations of total and free T PGSs on endpoints common to both sexes. Here, aligning with the results from the MR analyses, 15/20 of the nominally significant (p<0.05) associations remained similar (Z-test p>0.05 for difference in PGS effects) also in the cross-sex analyses,

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

pointing to shared genetic etiology rather than T action as the likely cause for these associations. For example, both female and male total T PGSs, with and without SHBG adjustment, associated with hypothyroidism risk with a similar effect size in the opposite sex, and all PGS associations to T2D were replicated in the other sex (Figure 4A, and Supplementary Tables 23&24).

The sex-specific associations were not replicated (Z-test p<0.05) in the opposite sex for total T PGS and stroke in both sexes, and the male free T PGS, head injury risk and osteoporosis (Figure 4A, and Supplementary Tables 23&24). In addition, the effect of male total T PGS on statin use was attenuated in females. Intriguingly, while both male and female total T PGSs associated with increased risk for stroke in females (HR=1.03, p=0.017 and HR=1.04, p=0.0079, respectively), the same PGSs rather associated with a reduced stroke risk in males (HR=0.97, p=0.017 and HR=0.98, p=0.073). This was the only endpoint for which we detected statistically significant evidence for a PGS having antagonistic effects depending on sex. The antagonistic effect of male total T PGS on stroke enhanced with SHBG adjustment (HR=0.96, p=0.011 in males and HR=1.07, p=0.0010 in females). In contrast, the female total T PGS showed no association with stroke in males after the SHBG adjustment (HR 0.99, p=0.65), suggesting all the PGS associations to stroke may not be fully androgen-driven. The results nevertheless indicate the genetic effects on stroke risk may be partly sex-specific (35), with potential interplay from sex hormones.

Finally, the cross-sex analyses further implied increased androgen load as a direct contributor to poorer reproductive health in females, agreeing with the MR-based causality assessments. For the reproductive endpoints that were associated with female free T PGS, i.e., infertility, PMB, PCOS and hirsutism, the male free T PGS had no predictive power (all p<0.05, Figure 4B, Supplementary Tables 23&24). Yet, this approach did not fully support exclusive sex-specific causality of free T for breast cancer in females and prostate cancer in males. Here, we observed that the effect sizes of the PGS associations were attenuated, but not in a statistically significant manner (p>0.05, Figure 4B). Echoing the findings from the causality analyses, these results thus suggest some degree of shared genetic risk, irrespective of T levels, between male and female hormonal cancers.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

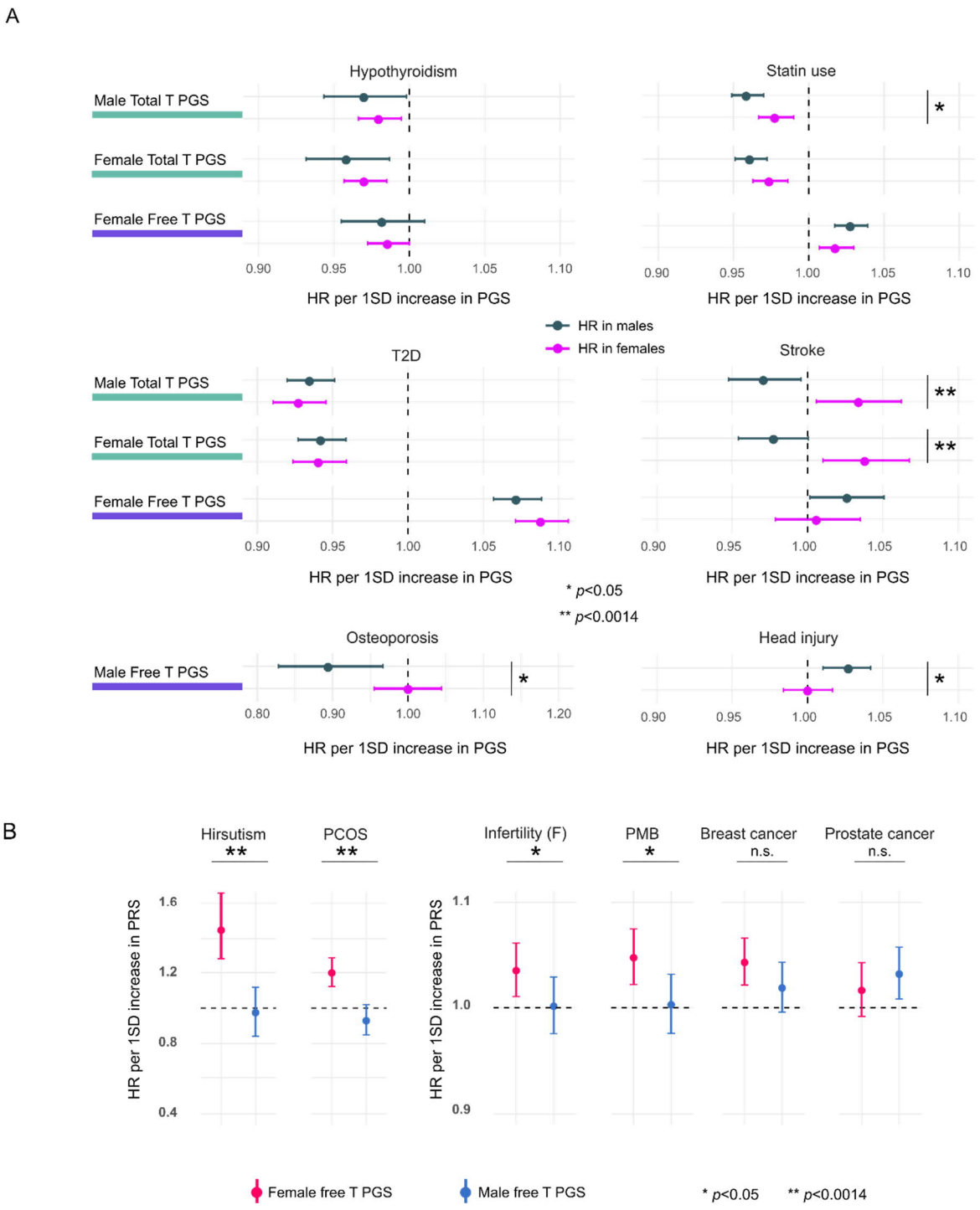


Figure 4. Results from the cross-sex analyses. A) Illustration of cross-sex PGS associations with hypothyroidism, statin use, T2D and stroke. $p<0.05$ suggests that the effects of a given PGS vary depending on sex B) Illustration of cases where only the female-specific free T PGS associates significantly with a trait and there is statistical evidence for male and female PGSs having different effects (hirsutism, polycystic ovary syndrome (PCOS), infertility and post-menopausal bleeding (PMB). However this did not held for breast and prostate cancers.*Chi-Squared $p<0.05$, ** = Chi-Squared $p<0.0014$, n.s. = not significant. Panels A and B show HR point estimates with 95% confidence intervals.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Extending the FinnGen discoveries

We next sought to validate and refine the FinnGen discoveries in additional datasets, extending our analysis to include quantitative traits not available in large numbers in FinnGen. To this end, we used genetic correlation analysis, allowing for estimation of the extent to which two traits are affected by the same genetic factors (36, 37), followed by causality estimations.

We selected 44 traits with publicly available GWAS summary statistics, identical to (e.g., T2D, breast and prostate cancers) or closely reflecting the studied disease phenotypes (heel bone mineral density (HBMD), mood swings) from FinnGen, adding anthropometric traits to the analyses (Supplementary Table 25). For most female-specific phenotypes studied in FinnGen, including hirsutism and PMB, we had no comparable phenotypes, as there are no published GWAS available.

We found evidence of significant genetic correlation in 72/352 instances (p<0.0011, corresponding to Bonferroni correction for 44 independent tests) (Figure 5, Supplementary Table 25) with the results reflecting the FinnGen PGS associations. We observed significant genetic correlations to traits related to metabolism, including many biomarkers and anthropometrics, but detected only few correlations to behavioral traits, and no significant correlations to neurological or temperamental traits (Figure 5, Supplementary Table 25). Notably, in all cases where we observed genetic correlation to behavioral traits, these had clear links to metabolism (smoking, sleep duration and exercise).

The genetic factors increasing serum total T and SHBG appeared to promote a favorable metabolic profile in males, supporting the PGS findings. Despite correlating with increased BMI, total T and SHBG were positively correlated to adiponectin, high-density lipoprotein (HDL) and lower waist-to-hip ratio (WHR) (r_g>0.20, p<0.0011) whilst lowering triglycerides and T2D incidence in males (r_g<-0.25, p<0.0011). In contrast, higher FAI and free T fractions in females associated with negative metabolic effects, including higher WHR (r_g=0.25, p=1.6e-22) and lower HDL cholesterol (r_g=-0.18, p=4.2e-05), whilst free T showed no significant correlations to these metabolism-related traits in males. Strong correlations were observed also between SHBG and

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

metabolic traits in females, including negative associations to markers of liver damage (alanine transaminase (ALAT), r_g=-0.21, *p*=1.7e-08) and gamma-glutamyl transferase (GGT), r_g=-0.16, *p*=0.0001).

We observed significant genetic correlations to hormonal cancers in both sexes, in line with recent findings (22, 25). In males, the genetic factors increasing FAI and free T promoted prostate cancer (r_g=0.12, *p*=0.0004), whereas in females these increased especially the risk of estrogen receptor (ER) positive breast cancer (r_g=0.15, *p*=9.1e-09). Finally, consistent with existing genetic, epidemiological and experimental data (1, 25), the genetic correlation analyses pointed towards shared genetic background for T, hemoglobin levels and body fat in both sexes (e.g. r_g=0.15, *p*=1.0e-07 and r_g=-0.14, *p*=0.0002 for hemoglobin and body fat, respectively, with male free T).

In the LCV and MR-Egger analyses we found statistically significant evidence (*p*<0.0011) of a causal relationship in 7% (26/354) of instances across the 44 traits (Figure 5 and Supplementary Table 26). Reflecting the results from FinnGen, despite the genetic correlations to many traits there was no evidence for significant causality, but the few suggested causal relationships involved traits with clear biological links to T function.

As examples of expected causal relationships (1, 25, 38), our analyses supported the contribution of free T levels to male-pattern baldness (MPB, GCP=0.44, *p*=7.7e-19) and hemoglobin levels (GCP=0.64, *p*=0.00085) (Figure 5 and Supplementary Table 26). In addition, despite the lack of significant genetic correlations between these traits, higher T in males was linked to increased number of children fathered (NCF, GCP=0.62, *p*=1.2e-15 for total T, GCP=0.42, *p*=0.011 for free T), raising speculation about the potential evolutionary benefits of males maintaining adequate T levels.

The MR Egger analyses supported the causality of female total T and free T to ER+ breast cancer (β=0.282, *p*=0.00014 and β=0.250, *p*=0.002, respectively). Instead of T, in these analyses prostate cancer risk was linked to SHBG levels in males (GCP=-0.64, *p*=0.00019) (39). Additionally, SHBG increased lymphocyte count in both sexes (GCP=0.39, *p*=1.6e-07 in males, GCP=0.42, *p*=0.00014 in females) (40), and in males, we linked SHBG also with reduced risk of erectile dysfunction (GCP=0.42, *p*=0.00032). In females, we observed causality

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

between SHBG and age at menopause (β=-0.756, *p*=0.0010), suggesting that inherited differences in SHBG levels likely modify reproductive phenotypes across female lifespan.

In most instances, we however found no evidence of a significant causal relationship between T levels and the studied traits (Supplementary Table 26). Yet, emphasizing the intricate relationship between hormone levels and metabolism (41), some connections to metabolism-related biomarkers emerged: for example, it appeared that triglycerides may causally influence T (GCP=-0.48, *p*=1.3e-30) and free T (GCP=-0.80, *p*=1.4e-06) levels in females. In combination with the PGS analyses, these results thus further suggest that whilst some sex-biased phenotypes may be directly related to T levels, in most instances T's relationship to complex traits and diseases is not straightforward.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

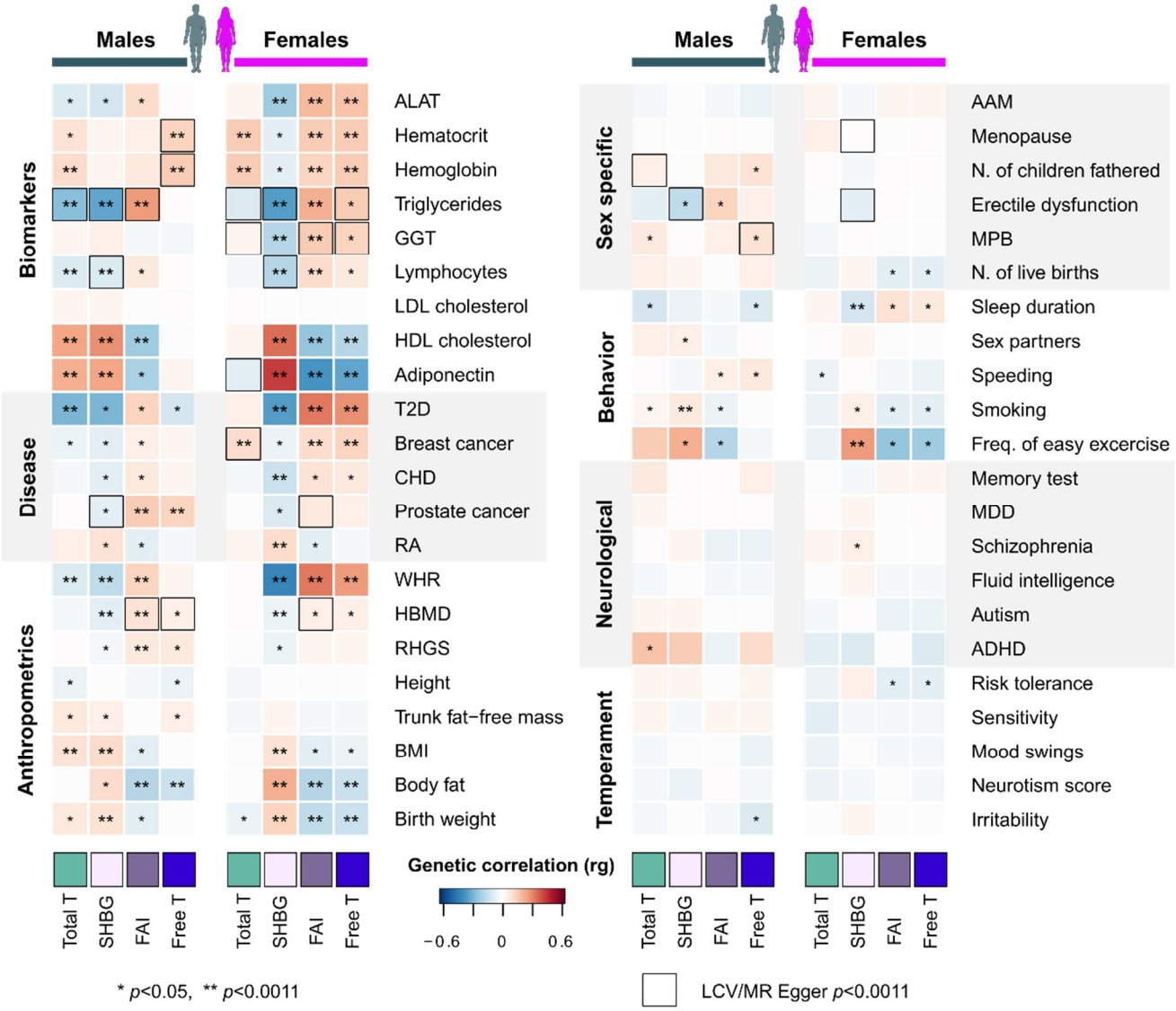


Figure 5. Genetic correlation and causal relationships between the 44 traits and phenotypes from public GWAS data based on LDSC and MR analyses. * $p<0.05$, ** $p<0.0011$ (significant correlation after Bonferroni correction for 44 independent tests). Black boxes = ($p<0.0011$) for a causal relationship between the tested traits based on LCV and MR Egger analyses. AAM = age at menarche, ALAT = alanine transaminase, ADHD = attention deficit and hyperactivity disorder, CHD = Coronary heart disease, GGT = Gamma-glutamyl transferase, HBMD = Heel bone mineral density, MDD = major depressive disorder, MPB = male pattern baldness, RA = rheumatoid arthritis, RHGS = relative handgrip strength, WHR = waist-to-hip ratio.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Discussion

Since its discovery in the early 20th century, testosterone (T) has been proposed to modify phenotypes and diseases that differ between the sexes, due to the extensive male-female differences in circulating T levels. To provide a broad and systematic perspective into the involvement of T as a regulator of health and disease in males and females, we have used genetic data as an anchor allowing us to address the potential causal role of post-pubertal T exposure for a wide range of diseases and traits thought to be under hormonal control.

We leveraged the UK Biobank resource to construct PGSs - predictors for genetically determined T levels - for both males and females, which we then associated with a uniquely rich collection of disease endpoints from 217,464 participants in FinnGen. This combination allowed us to extend and refine recent efforts that have utilized genetic data to understand the disease impacts of T, concentrating on a limited set of phenotypes or only to males (22, 23, 25-27). For instance, we included many novel endpoints from FinnGen, including registry-based phenotypes like statin use, female-specific endpoints such as hirsutism and post-menopausal bleeding, and neurological diagnoses like conduct disorder, where hormonal contribution has not been previously studied using genetic approaches. In addition, by taking advantage of the sex-specific genetic determinants of T in cross-sex analyses, and through careful Mendelian Randomization strategies, we could pinpoint T's causal effects on adult health.

Post-pubertal T exposure and disease risk

Based on our analyses, three major themes emerged regarding T's contribution to disease. First, contrary to T's established role as a male androgen, we report that the studied PGSs associated with disease risk especially in females. Secondly, we highlight distinct association profiles for total T and free T in both sexes, consistent with proposed divergent biological effects for the bound and unbound T fractions (13-15). Underscoring the potential role of SHBG as a confounder when studying the action of T, the former closely correlates with SHBG levels in males, and the latter in females (13-15). Thirdly, in many, but not all instances, the associations with the T and free T PGSs truly seemed to reflect androgen action. We observed causal relationships between a genetic predisposition to higher free T levels and several sex-specific and sex-biased

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

phenotypes with clear biological links to T, but less contribution to most other phenotypes, echoing experimental data and findings from recent MR studies (22, 25).

Overall, our study highlights T's intricate connections to reproduction and metabolism. Besides the causal links to several female-specific reproductive endpoints, many of the significant associations involved diseases and traits from the metabolic and endocrine categories. We stress the complexity of these relationships, as for many metabolic traits SHBG – either directly or potentially through its action on other hormones - appeared to confound T's associations and causality estimates. Beyond this T/SHBG relationship, we speculate that some of the observed associations to metabolic health, including T2D, may reflect even more widespread genetic pleiotropy and thus overall complex shared genetic etiology rather than T action.

Due to the lack of significant associations and causality for many other traits, we anyhow speculate that normal variation in T levels - contrary to popular beliefs - has only modest effects on most phenotypes. Particularly, the grounds to explain some temperamental and neurological phenotypes like anxiety and emotional instability with heritable differences in adult T levels (42, 43) appears unsubstantiated in the light of our study. Although we cannot exclude a causal relationship between T and some of these phenotypes, our data suggests that especially without larger sample sizes or refined phenotyping, efforts to establish relationships between T levels and e.g. behavior will likely be unproductive. Taken together, supporting recent recommendations, our data thus suggests that for example the risks and benefits of using T as a medical treatment should be carefully weighted, given T's complex and indirect relationship to most phenotypes and potential adverse and beneficial outcomes in both sexes (7, 25).

T as a contributor to sex differences

Having comprehensively mapped the impacts of T across diverse complex disease and traits, we can start drawing inferences on the role of T as a contributor to the male-female differences. We reason that should T causally contribute to within-sex variability of a trait, T should similarly be accountable for a fraction of the between-sex difference detected in the same trait. Indeed, in such instances where causality of T was implicated for a sex-shared trait, the effect estimates often aligned with the direction of the phenotypic sex

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

bias in the given trait, i.e., higher T levels associated with typical male characteristics. For instance, our work highlighted causal connections between increasing free T, higher hemoglobin and higher bone strength, backed up by previous experimental observations (1, 34, 44) and thus suggests the involvement of T as a mediator of the established sex differences in these characteristics. Extrapolating from the MR estimates, we estimate that ~10-20% of the mean difference in hemoglobin levels between males and females may result from average differences in free T levels, consistent with the notion that T directly affects male-female differences in, e.g., athletic capacity (1). Moreover, we find that higher free T levels are causal to masculine external features like hirsutism in females and baldness in males, further implicating direct involvement of T in defining sex differences. Additionally, in females higher bioavailable T levels correlated genetically with a shift of metabolism into a male-like direction (e.g., increased BMI, WHR, and poorer blood lipids, yet reduced risk for obesity and reduced body fat levels).

Overall, the limited evidence for the causal involvement of T for most of the studied traits suggests that most phenotypic sex differences are not attributable to a linear relationship between T levels and a phenotype. Instead the impact of T can be mediated through a threshold effect, potentially at a given developmental time window, acting as a switch that results in more global rewiring of biological processes and thereby in systematic male-female differences. Under this model, the within-sex variability in T may have non-existent or very subtle effects on many phenotypes, complicating the detection of potential causality of T in a study setting like ours that utilizes normal physiological ranges of adult T levels.

Connection between reproductive success and distinct genetic architecture of T between sexes?

Finally, our study provides some unique insight into the potential causes behind the sex differences in T levels. In the cross-sex PGS analyses, we found only one case where there was clear evidence of antagonistic effects for the T PGS between the sexes, which in theory might promote sex-differences in T levels. Both male- and female-specific total T PGS protected from stroke in males, whereas the same PGSs had exactly the opposite effects on stroke risk in females. Although we cannot conclude that the action of T truly drives

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

all of these associations, the result agrees with a degree of sex-specificity in the genetic disease mechanisms for stroke (35).

The observation that the genetic variants responsible for regulating T levels are largely distinct between males and females anyhow raises speculation about the evolutionary forces maintaining sex differences and shaping the biology of T. The cross-sex genetic correlations for traits related to fitness (e.g., reproductive success) are generally expected to be low, due to potentially conflicting evolutionary pressures (45). We indeed associated T positively with reproductive success in males (T increasing the number of children fathered), but negatively with both pre- and postmenopausal reproductive health in females with evidence for a causal role of T behind these associations. We may thus speculate that there exists a selective advantage in relation to reproductive success to maintain higher T levels in males, whereas the opposite may be true for females, potentially promoting the widespread sex differences across multiple traits.

Challenges for genetic analyses in addressing T function

In this study, the combination of data from two independent biobanks should result in reduced confounding, allowing robust inferences about the contribution of T to human phenotypes. Although based on extensive data sets, our study still has some limitations. We stress that studying T levels differs drastically from studying T action, the serum T levels serving only as a proxy for the latter. Theoretically, variants increasing T activity may include 1) variants increasing T production or decreasing T metabolism and breakdown, resulting in consistently higher T levels and 2) variants increasing T uptake in peripheral tissues or the sensitivity of body to T, lowering circulating T levels. This severely limits the potential of standard genetic methods, expecting linear relationships between two endpoints, to find links between T levels and complex traits. In light of the established model of regulation of T levels via the hypothalamic-pituitary-gonadal (HPG) axis, (high T levels leading to downregulation of gonadotropin secretion, ultimately lowering T levels), this limitation may apply especially to studying T action in males (16, 17).

Moreover, likewise to other recent studies (22, 25), we emphasize the challenges the close ties between SHBG and T may pose. Earlier studies have addressed this issue by attempting to distinguish genetic variants

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

with more specific effects for testosterone by using clustering-based approaches (22, 25). We, besides accounting for the effect of SHBG by calculating the level of bioavailable T, additionally examined how SHBG-adjustment influences the results of the testosterone PGS analyses and causality estimates, allowing for the detection of some traits with clear SHBG confounding.

Besides this specific example, genetic studies are also prone to confounding by pleiotropy in a more general sense, whereby a gene influences multiple traits via independent biological pathways (30). Although we opted for adjusting for the effect of BMI in our GWAS – obesity being a known confounder for testosterone levels (Supplementary Figure 8) (46, 47) - it remains possible that this connection still affects our findings. Genetic pleiotropy may confound also MR analyses, and despite the vast potential of MR in establishing causal relationships (48), we generally propose caution in interpreting these findings. In our case all MR models did not always agree on causality (Supplementary Tables 22 and 27), common to recent MR studies assessing the function of T (22, 23, 25). When based on large number of variants, the MR analyses often include variants that do not fulfill the strict definition of an instrumental variable. Here we show via MR and cross-sex PGS analyses that such pleiotropic variants may underlie for example the T associations with T2D and hypothyroidism.

Finally, our setting does not allow for assessing the effects of fetal T exposure, which may be critical, e.g., for neurological traits (49). We also emphasize that our results are based on normal variation in T levels, not on the effects of supraphysiological T injections. Additionally, many of T's effects depend on its conversion to estradiol also in males, and we cannot rule this out as a potential confounder in our study. Finally, the data used in our study does not allow for assessing the effects of acute changes in hormone secretion, and personal differences in the response to such fluctuations may be crucial for some phenotypes.

Despite these challenges, we were able to highlight several novel albeit often expected relationships with genetically determined T levels, human health, and sex differences. Overall, we show the power of biobank-scale genetic analyses to extend and clarify the results obtained from epidemiological and experimental studies, leading to improved understanding on how human phenotypes are related to sex hormone levels.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Conclusion

Here, by combining PGS and electronic health records, we have assessed the link between T, disease risk and complex traits, including many understudied phenotypes. We provide a broad genetic perspective on disease impacts and phenotypic effects of post-pubertal T exposure, extending previous experimental and MR studies. We shed light on the interplay between T and complex traits in both sexes, and the role of T in driving sex differences. Finally, we underscore some critical factors that should be taken into account when assessing these relationships, providing a reference point for future genetic and epidemiological studies studying the action of T.

Methods

Genotype and phenotype data from the UK Biobank

The genetic association analysis was based on data from the UK Biobank, a population-based biobank consisting of 502 637 subjects (aged 37-73 years)(50). At recruitment, participants provided electronic signed consent. Ethics approval for the UK Biobank study was obtained from the North West Centre for Research Ethics Committee (11/NW/0382). All experiments were performed in accordance to relevant guidelines and regulations including the Declaration of Helsinki ethical principles for medical research. This study was run under UK Biobank application number 22627.

Multiple biochemical assays have been performed on the entire UK Biobank cohort, including measurements of serum testosterone, SHBG and serum albumin. These measurements were performed once for each participant for using Beckman Coulter DXI 800 Chemiluminescent Immunoassay, with competitive binding for testosterone and two-step sandwich for SHBG, referenced in https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf. The document contains also details about the extensive QC procedures for all biochemical measurements in the UK Biobank. For example, assay consistency was monitored by using internal QC samples between batches, and external quality assurance schemes against the ISO 17025:2005 standard.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We restricted our study to encompass 408,186 individuals from the white British subset. We removed outliers for genotype heterozygosity and missingness, as well as samples with sex chromosome aneuploidies, mismatches between reported and inferred sex, and samples that UK Biobank did not use in relatedness calculations (50). We did not exclude related samples since our analysis method (BOLT-LMM) allows for their inclusion.

For the GWASs, we used biochemically measured testosterone and SHBG, and calculated FAI and free T. For calculation of FAI, we used the formula *100*Testosterone/SHBG* (nmol/ml). Calculated free T was derived using the Vermeulen equation using directly measured albumin values for each participant in the equation, as described in (51, 52).

All four traits were separated by sex and log transformed. A linear regression model was fit for each trait with BMI and age as covariates, as well as menopause status for females. Subjects with residuals values of +-5 SD from the mean were excluded from the analyses, serving as a further QC step to exclude outliers potentially reflecting medical conditions or drug use affecting androgen levels. Inverse normalized values of the remaining residuals were used as phenotype values for the GWAS analyses. After the QC-steps, our study included altogether 177,499 males and 205,141 females.

Genetic association analysis and definition of the lead SNPs

The GWAS analyses were performed using BOLT-LMM (v2.3.2)(51, 52). Imputed SNPs were restricted to variants with MAF ≥ 0.1 % and imputation quality ≥ 0.7 (50). 1000 Genomes European data was used as reference LD scores for calibrating the BOLT-LMM statistic. First 10 Principal Components were used as quantitative covariates in the runs. A linear regression model was fit for each trait with BMI and age as covariates, as well as menopause status for females. Genetic correlation analyses, heritability estimates and number of loci found implied the results remained consistent with different covariate configurations (Supplementary Figure 8), but we chose to include body mass index (BMI), known to associate with T levels (28, 29), as a covariate in our GWAS. Including up to 127 covariates (based on (53)), e.g., assay center, dilution

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

factors, blood draw time, and socioeconomic status indicators, or excluding related individuals from the analysis all showed negligible effects on the genetic findings we report here (Supplementary Figure 8).

Independent lead SNPs were selected for each chromosome by recursively taking the SNP with the lowest p-value (until none below the *p*-value threshold 5e-08 were left) from the GWAS summary statistics and removing all SNPs 500kb on each side of it from the next round. The chromosomal positions of these 1mb windows were stored, and overlapping windows were merged into the final list of loci. The SNP with the lowest p-value in each of these windows was selected as the lead SNP.

Pathway, tissue enrichment and co-localisation analyses

Tissue and gene set enrichment analyses were carried out with SNP2GENE and GENE2FUNC implemented in FUMA using default settings (54). For testing in which tissues the genes residing in the GWAS loci were preferably expressed, we used the full distribution of SNP p-values and the GTEx v6 30 general tissue types. For pathway analysis, to assess whether the genes in the GWAS loci are overrepresented in pre-defined gene sets via hypergeometric tests, we selected manually curated KEGG-pathways (55). For co-localisation analyses to assess whether the genetic loci showed evidence for shared genetic effects between males and females, and to estimate the maximum posterior probability (MAP) for the loci being shared, we used gwas-pw (56).

Replication in the Young Finns Cohort and calculation of PGS

The Cardiovascular Risk in Young Finns Study (YFS) is a longitudinal follow-up of 3,596 subjects at baseline. The baseline survey was conducted in 1980 and subsequent follow-ups involving the whole sample were held in 1983, 1986, 2001, 2007, 2011, and 2017. For testosterone and SHBG, we used data on 2001 follow-up (Subjects aged 24-39 yrs). A venous blood sample was drawn from antecubital vein after 12-hour overnight fast. Serum was aliquoted and stored in -70 Celsius degrees until analysis. In males, total testosterone quantification was performed in 2009 with competitive radioimmunoassay (Spectria Testosterone kit, Orion Diagnostica, Espoo, Finland) and Bio-Rad Lyphocheck control serums 1, 2, and 3 were used in quality control. Before quantification, serum aliquots had been melted three times. Total testosterone was quantified first

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

and aliquots were re-frozen before SHBG quantification. In females, total testosterone quantification was performed in 2011. SHBG quantification for males was done in 2009 and for females in 2011 with Spectria SHBG IRMA kit (Orion Diagnostica, Espoo, Finland). Free testosterone was estimated using Vermeulen's formula. As albumin concentration was not available, we used fixed albumin concentration of 43 g/l.

In YFS, genotyping was performed at the Wellcome Trust Sanger Institute (UK) using customized Illumina Human Map 670k bead array. The custom content on 670k array replaced some poor performing probes on Human610 and added more CNV content. As quality control, we excluded individuals and probes with over 5% of missingness (--geno and --mind filters in Plink). Variants deviating from Hardy-Weinberg equilibrium (p < 1x10^-6) and minor allele frequency below 1% were excluded. Related samples were excluded (n=51) with pi-hat cut-off of 0.2. Total of 2,442 individuals and 546,674 variants passed the quality control measures. The mean call rate across all included markers after the quality control was 0.9984. Next, imputation was performed using population-specific Sequencing Initiative Suomi (SISu) as reference panel. We examined the association of the UK Biobank GWAS lead SNPs with the corresponding T trait in YFS. If the annotated lead SNP was not available in YFS, we used LDstore2 (v2.0b) (57) to calculate LD in a 100kb window around the lead SNP in the UK Biobank imputed data and selected the closest SNP with R2>0.8 with the lead SNP as a proxy.

To construct PGS we applied the LDpred (28) method to the sex-specific GWAS results from the UK Biobank for total T, SHBG, FAI and free T using 1000 Genomes Europeans as LD reference and the default LD radius to account for LD. We then used the weights from the LDpred infinitesimal model to construct genome-wide PGSs for each individual in the YFS with Plink 2.0 (11 Feb 2018). Only variants imputed with high confidence (imputation INFO $\geq$ 0.8) were included in PGS calculation. Variants in chromosomes 1-22 and chrX were included. In males, allele dosage of 2 was used for X-chromosomal haploid variants. To evaluate the PGS prediction accuracy in the YFS, we calculated the R2 for each trait using using linear regression with z-score normalized PGS as predictor, age and 10 PCs as covariates and z-score normalized T trait as outcome.

Estimation of heritability and calculation of the genetic correlations by LDSC

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

SNP-based heritability for the studied T traits and genetic correlations between these and with 44 additional phenotypes were estimated using linkage disequilibrium score regression (LDSC)(36). The summary statistics for the 44 traits were downloaded directly from the source repositories and analysed locally, for the original sources please see references in Supplementary Table 25. For the genetic correlation analyses, pre-computed LD Scores from 1000 Genomes Europeans excluding the HLA region were used. For 23 traits we performed the analyses using sex-specific GWAS results and compared these to data from sex-combined GWAS (Supplementary Table 27). Generally, genetic correlation results using either sex-specific or sex-combined GWAS data were highly similar.

Disease associations in FinnGen

To assess if the PGS for studied traits associate with disease risk we utilized the FinnGen study (data freeze 5), consisting of 217,464 (94,478 males, 122,986 females) (29). FinnGen is comprised of Finnish prospective epidemiological and disease-based cohorts and voluntary biobank samples collected by hospital biobanks. The genotypes have been linked to national hospital discharge (available from 1968), death (1969–), cancer (1953–) and medication reimbursement (1964–) registries as well as the registry on medication purchases (1995-). The samples were genotyped with Illumina and Affymetrix arrays (Illumina Inc., San Diego, and Thermo Fisher Scientific, Santa Clara, CA, USA). The genotypes have been imputed with using the SISu v3 population-specific reference panel developed from high-quality data for 3,775 high-coverage (25-30x) whole-genome sequencing in Finns. The detailed genotype imputation workflow can be found at https://dx.doi.org/10.17504/protocols.io.xbgfijw. The dataset uses genome build 38 (hg38).

For PGS analyses, we used same variant weights (LDpred infinitesimal model) as for YFS, and calculated genome-wide PGSs for each individual with PLINK2 (v2.00a2.3LM). Variants in chromosomes 1-22 and chromosome X (imputed with high confidence, imputation INFO ≥0.7) were included (total number of variants ranging from 6,535,263 for female total T to 6,536,405 for female SHBG) and we used genotype dosages to incorporate imputation uncertainty. In males, allele dosage of 2 was used for X-chromosomal haploid variants. We studied the PGS associations to 36 disease endpoints with potential links to androgens,

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

representing six loosely defined disease categories. For details of the studied phenotypes see Supplementary Table 21, www.finngen.fi and risteys.finngen.fi. Cox proportional hazards models were used for estimating hazard ratios (HRs) and 95% CIs, with age as the time scale and 10 first principal components of ancestry and genotyping batch as covariates. The proportionality assumption for Cox models was assessed with Schoenfeld residuals and log-log plots. For the cross-sex analyses, we took the sex-specific PGSs, and checked whether these would associate with the studied endpoints in the other sex, using the z-test to compare equality between the original and cross-sex associations. We additionally performed SHBG-adjusted PGS associations to all endpoints to control for potential confounding of SHBG to total and free T.

Patients and control participants in FinnGen provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, older research cohorts, collected prior the start of FinnGen (in August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after approval by Valvira, the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Valvira. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) approved the FinnGen study protocol Nr HUS/990/2017.

The FinnGen study is approved by Finnish Institute for Health and Welfare (THL), approval number THL/2031/6.02.00/2017, amendments THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019, Digital and population data service agency VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3 the Social Insurance Institution (KELA) KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, and Statistics Finland TK-53-1041-17. The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 6 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, Auria Biobank AB17-5154, Biobank Borealis of Northern Finland_2017_1013, Biobank of Eastern Finland 1186/2018, Finnish Clinical Biobank Tampere MH0004, Central Finland Biobank 1-2017, and Terveystalo Biobank STB 2018001.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Causality analyses

MR analyses treat genetic variants as instrumental variables and their reliability depends on two key assumptions: 1) alleles are randomly assigned, and 2) that alleles that influence exposure do not influence the outcome via any other means. The first assumption is controlled by using BOLT-LMM as our model in the primary GWAS analysis, but the second is harder to control when using a large number of SNPs as instrumental variables. In line with the observed wide-spread genetic pleiotropy affecting most complex traits, we noted that the GWAS loci contained many genes associated with pleiotropic effects on human phenotypes (for example, *LIN28B (58), GCKR (56)* and *TYK2(59)*). Therefore, given the vast polygenicity of the studied T traits, we chose latent causal variable (LCV) (31) and MR-Egger (32) as our primary MR methods, designed to take into account pleiotropy-induced confounding when assessing causal relationships. LCV has been proposed to provide more unbiased causality estimates than conventional MR approaches, whereas MR Egger should provide accurate causality estimates under the InSIDE assumption (the genetic variants have pleiotropic effects that are independent in magnitude and are thus not mediated by a single confounder exposure), besides its recommended use as a sensitivity analysis for conventional MR (31, 32). For comparison we also ran conventional MR analyses (Inverse-Variance Weighed (IVW)), that remains more sensitive for confounding by genetic correlation and pleiotropy (31). To extend the basic MR Egger analysis and to tease out the potential effects of SHBG on causality estimates of total and free T, we used multivariable MR Egger (33). LCV reports genetic causality proportion (GCP) as an estimate of causality, under a model where genetic correlation between two traits is mediated by a latent variable having a causal effect on each trait. GCP=1 means trait 1 is fully correlated with the latent variable, and hence fully causal to trait 2. A high GCP value and a statistically significant effect support partial genetic causality between the traits, and suggest that interventions targeting trait 1 are likely to affect trait 2. The *p*-value obtained in the analysis refers to the null hypothesis that the GCP=0. A highly significant *p*-value does not require a high GCP. Positive GCP value indicates causality of trait1 to trait2, whereas a negative value indicates support for causality of trait 2 to trait 1. LCV also estimates genetic correlation between the traits. To estimate the how much T could explain sex differences in hemoglobin we calculated ((FT_m-FT_f/SD_FT_m)*β*SD_H_m)/((H_m-

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

H_f)/SD_H_m), where FT = mean free Testosterone, m=males, f=females, SD = standard deviation, β=MR estimate, H=Hemoglobin in UK Biobank based on (60). We applied the LCV, MR-Egger, multivariate MR-Egger and IWV models locally using R 4.0.2. The MR analyses were run using TwoSampleMR (v0.5.2) (61) and MendelianRandomization (v0.5.0) R packages (62). For the traits from public GWAS included in genetic correlation analyses, in 16 out 44 instances we could perform two sample MR (phenotype data not based on UK Biobank samples, Supplementary Table 24). FinnGen represents an independent research cohort from the UK Biobank and thus all FinnGen causality analyses were two-sample MR analyses.

References

1. Handelsman DJ, Hirschberg AL, Bermon S. Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance. Endocr Rev. 2018;39(5):803-29.
2. Tyagi V, Scordo M, Yoon RS, Liporace FA, Greene LW. Revisiting the role of testosterone: Are we missing something? Rev Urol. 2017;19(1):16-24.
3. Davis SR, Wahlin-Jacobsen S. Testosterone in women—the clinical significance. The Lancet Diabetes & Endocrinology. 2015;3(12):980-92.
4. Vandenput L, Ohlsson C. Genome-wide association studies on serum sex steroid levels. Mol Cell Endocrinol. 2014;382(1):758-66.
5. Kloner RA, Carson C, Dobs A, Kopecky S, Mohler ER. Testosterone and Cardiovascular Disease. J Am Coll Cardiol. 2016;67(5):545-57.
6. Gagliano-Jucá T, Basaria S. Testosterone replacement therapy and cardiovascular risk. Nature Reviews Cardiology. 2019;16(9):555-74.
7. Bhasin S, Brito JP, Cunningham GR, Hayes FJ, Hodis HN, Matsumoto AM, Snyder PJ, Swerdloff RS, Wu FC, Yialamas MA. Testosterone Therapy in Men With Hypogonadism: An Endocrine Society* Clinical Practice Guideline. The Journal of Clinical Endocrinology & Metabolism. 2018;103(5):1715-44.
8. Celec P, Ostatníková D, Hodosy J. On the effects of testosterone on brain behavioral functions. Front Neurosci. 2015;9:12-.
9. Venturelli E, Orenti A, Fabricio ASC, Garrone G, Agresti R, Paolini B, Bonini C, Gion M, Berrino F, Desmedt C, Coradini D, Biganzoli E. Observational study on the prognostic value of testosterone and adiposity in postmenopausal estrogen receptor positive breast cancer patients. BMC Cancer. 2018;18(1):651-.
10. Stattin P, Lumme S, Tenkanen L, Alfthan H, Jellum E, Hallmans G, Thoresen S, Hakulinen T, Luostarinen T, Lehtinen M, Dillner J, Stenman U-H, Hakama M. High levels of circulating testosterone are not associated with increased prostate cancer risk: A pooled prospective study. 2004;108(3):418-24.
11. Eisenegger C, Naef M, Snozzi R, Heinrichs M, Fehr E. Prejudice and truth about the effect of testosterone on human bargaining behaviour. Nature. 2010;463(7279):356-9.
12. Clark RV, Wald JA, Swerdloff RS, Wang C, Wu FCW, Bowers LD, Matsumoto AM. Large divergence in testosterone concentrations between men and women: Frame of reference for elite athletes in sex-specific competition in sports, a narrative review. 2019;90(1):15-22.
13. Goldman AL, Bhasin S, Wu FCW, Krishna M, Matsumoto AM, Jasuja R. A Reappraisal of Testosterone's Binding in Circulation: Physiological and Clinical Implications. Endocr Rev. 2017;38(4):302-24.
14. Cunningham SK, Loughlin T, Culliton M, McKenna TJ. The Relationship between Sex Steroids and Sex-Hormone-Binding Globulin in Plasma in Physiological and Pathological Conditions. 1985;22(5):489-97.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

15. Mendel CM. The Free Hormone Hypothesis: A Physiologically Based Mathematical Model*. Endocr Rev. 1989;10(3):232-74.
16. Pasquali R. The hypothalamic-pituitary-adrenal axis and sex hormones in chronic stress and obesity: pathophysiological and clinical aspects. Ann N Y Acad Sci. 2012;1264(1):20-35.
17. Trost LW, Mulhall JP. Challenges in Testosterone Measurement, Data Interpretation, and Methodological Appraisal of Interventional Trials. The journal of sexual medicine. 2016;13(7):1029-46.
18. Ohlsson C, Wallaschofski H, Lunetta KL, Stolk L, Perry JRB, Koster A, Petersen A-K, Eriksson J, Lehtimäki T, Huhtaniemi IT, Hammond GL, Maggio M, Coviello AD, Group ES, Ferrucci L, Heier M, Hofman A, Holliday KL, Jansson J-O, Kähönen M, Karasik D, Karlsson MK, Kiel DP, Liu Y, Ljunggren Ö, Lorentzon M, Lyytikäinen L-P, Meitinger T, Mellström D, Melzer D, Miljkovic I, Nauck M, Nilsson M, Penninx B, Pye SR, Vasan RS, Reincke M, Rivadeneira F, Tajar A, Teumer A, Uitterlinden AG, Ulloor J, Viikari J, Völker U, Völzke H, Wichmann HE, Wu T-S, Zhuang WV, Ziv E, Wu FCW, Raitakari O, Eriksson A, Bidlingmaier M, Harris TB, Murray A, de Jong FH, Murabito JM, Bhasin S, Vandenput L, Haring R. Genetic Determinants of Serum Testosterone Concentrations in Men. PLOS Genetics. 2011;7(10):e1002313.
19. Harris JA, Vernon PA, Boomsma DIJBG. The Heritability of Testosterone: A Study of Dutch Adolescent Twins and Their Parents. 1998;28(3):165-71.
20. Ruth KS, Campbell PJ, Chew S, Lim EM, Hadlow N, Stuckey BGA, Brown SJ, Feenstra B, Joseph J, Surdulescu GL, Zheng HF, Richards JB, Murray A, Spector TD, Wilson SG, Perry JRB. Genome-wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. Eur J Hum Genet. 2016;24(2):284-90.
21. Koenis MMG, Brouwer RM, van Baal GCM, van Soelen ILC, Peper JS, van Leeuwen M, Delemarre-van de Waal HA, Boomsma DI, Hulshoff Pol HE. Longitudinal Study of Hormonal and Physical Development in Young Twins. The Journal of Clinical Endocrinology & Metabolism. 2013;98(3):E518-E27.
22. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, Beaumont RN, Wittemans L, Martin S, Busch AS, Erzurumluoglu AM, Hollis B, O'Mara TA, McCarthy MI, Langenberg C, Easton DF, Wareham NJ, Burgess S, Murray A, Ong KK, Frayling TM, Perry JRB, The Endometrial Cancer Association C. Using human genetics to understand the disease impacts of testosterone in men and women. Nat Med. 2020;26(2):252-8.
23. Flynn E, Tanigawa Y, Rodriguez F, Altman RB, Sinnott-Armstrong N, Rivas MA. Sex-specific genetic effects across biomarkers. Eur J Hum Genet. 2020.
24. Sinnott-Armstrong N, Naqvi S, Rivas M, Pritchard JK. GWAS of three molecular traits highlights core genes and pathways alongside a highly polygenic background. eLife. 2021;10:e58615.
25. Mohammadi-Shemirani P, Chong M, Pigeyre M, Morton RW, Gerstein HC, Paré G. Effects of lifelong testosterone exposure on health and disease using Mendelian randomization. eLife. 2020;9:e58914.
26. Hayes BL, Robinson T, Kar SP, Ruth KS, Tsilidis KK, Frayling TM, Murray A, Martin RM, Lawlor DA, Richmond RC. Do sex hormones confound or mediate the effect of chronotype on breast and prostate cancer? A Mendelian randomization study. 2021:2021.04.20.21255783.
27. Harrison S, Davies NM, Howe LD, Hughes A. Testosterone and socioeconomic position: Mendelian Randomization in 306,248 men and women participants of UK Biobank. 2020:2020.11.06.20226944.
28. Vilhjálmsson Bjarni J, Yang J, Finucane Hilary K, Gusev A, Lindström S, Ripke S, Genovese G, Loh P-R, Bhatia G, Do R, Hayeck T, Won H-H, Ripke S, Neale Benjamin M, Corvin A, Walters James TR, Farh K-H, Holmans Peter A, Lee P, Bulik-Sullivan B, Collier David A, Huang H, Pers Tune H, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu Silviu A, Begemann M, Belliveau Richard A, Jr., Bene J, Bergen Sarah E, Bevilacqua E, Bigdeli Tim B, Black Donald W, Bruggeman R, Buccola Nancy G, Buckner Randy L, Byerley W, Cahn W, Cai G, Campion D, Cantor Rita M, Carr Vaughan J, Carrera N, Catts Stanley V, Chambert Kimberly D, Chan Raymond CK, Chen Ronald YL, Chen Eric YH, Cheng W, Cheung Eric FC, Chong Siow A, Cloninger CR, Cohen D, Cohen N, Cormican P, Craddock N, Crowley James J, Curtis D, Davidson M, Davis Kenneth L, Degenhardt F, Del Favero J, DeLisi Lynn E, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Essioux L, Fanous Ayman H, Farrell Martilias S, Frank J, Franke L, Freedman R, Freimer Nelson B, Friedl M, Friedman Joseph I, Fromer M, Genovese G, Georgieva L, Gershon Elliot S, Giegling I, Giusti-Rodrguez P, Godard S, Goldstein Jacqueline I, Golimbet V, Gopal S, Gratten J, Grove J, de Haan L, Hammer C, Hamshere Marian L, Hansen M, Hansen T, Haroutunian V, Hartmann Annette M, Henskens Frans A, Herms S, Hirschhorn Joel N, Hoffmann P, Hofman

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

A, Hollegaard Mads V, Hougaard David M, Ikeda M, Joa I, Julia A, Kahn Rene S, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller Matthew C, Kelly Brian J, Kennedy James L, Khrunin A, Kim Y, Klovins J, Knowles James A, Konte B, Kucinskas V, Kucinskiene Zita A, Kuzelova-Ptackova H, Kahler Anna K, Laurent C, Keong Jimmy Lee C, Lee SH, Legge Sophie E, Lerer B, Li M, Li T, Liang K-Y, Lieberman J, Limborska S, Loughland Carmel M, Lubinski J, Lnnqvist J, Macek M, Jr., Magnusson Patrik KE, Maher Brion S, Maier W, Mallet J, Marsal S, Mattheisen M, Mattingsdal M, McCarley Robert W, McDonald C, McIntosh Andrew M, Meier S, Meijer Carin J, Melegh B, Melle I, Mesholam-Gately Raquelle I, Metspalu A, Michie Patricia T, Milani L, Milanova V, Mokrab Y, Morris Derek W, Mors O, Mortensen Preben B, Murphy Kieran C, Murray Robin M, Myin-Germeys I, Mller-Myhsok B, Nelis M, Nenadic I, Nertney Deborah A, Nestadt G, Nicodemus Kristin K, Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh S-Y, Olincy A, Olsen L, Van Os J, Pantelis C, Papadimitriou George N, Papiol S, Parkhomenko E, Pato Michele T, Paunio T, Pejovic-Milovancevic M, Perkins Diana O, Pietilinen O, Pimm J, Pocklington Andrew J, Powell J, Price A, Pulver Ann E, Purcell Shaun M, Quested D, Rasmussen Henrik B, Reichenberg A, Reimers Mark A, Richards Alexander L, Roffman Joshua L, Roussos P, Ruderfer Douglas M, Salomaa V, Sanders Alan R, Schall U, Schubert Christian R, Schulze Thomas G, Schwab Sibylle G, Scolnick Edward M, Scott Rodney J, Seidman Larry J, Shi J, Sigurdsson E, Silagadze T, Silverman Jeremy M, Sim K, Slominsky P, Smoller Jordan W, So H-C, Spencer Chris CA, Stahl Eli A, Stefansson H, Steinberg S, Stogmann E, Straub Richard E, Strengman E, Strohmaier J, Stroup TS, Subramaniam M, Suvisaari J, Svrakic Dragan M, Szatkiewicz Jin P, Sderman E, Thirumalai S, Toncheva D, Tooney Paul A, Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb Bradley T, Weiser M, Wildenauer Dieter B, Williams Nigel M, Williams S, Witt Stephanie H, Wolen Aaron R, Wong Emily HM, Wormley Brandon K, Wu Jing Q, Xi Hualin S, Zai Clement C, Zheng X, Zimprich F, Wray Naomi R, Stefansson K, Visscher Peter M, Adolfsson R, Andreassen Ole A, Blackwood Douglas HR, Bramon E, Buxbaum Joseph D, Børglum Anders D, Cichon S, Darvasi A, Domenici E, Ehrenreich H, Esko T, Gejman Pablo V, Gill M, Gurling H, Hultman Christina M, Iwata N, Jablensky Assen V, Jonsson Erik G, Kendler Kenneth S, Kirov G, Knight J, Lencz T, Levinson Douglas F, Li Qingqin S, Liu J, Malhotra Anil K, McCarroll Steven A, McQuillin A, Moran Jennifer L, Mortensen Preben B, Mowry Bryan J, Nthen Markus M, Ophoff Roel A, Owen Michael J, Palotie A, Pato Carlos N, Petryshen Tracey L, Posthuma D, Rietschel M, Riley Brien P, Rujescu D, Sham Pak C, Sklar P, St. Clair D, Weinberger Daniel R, Wendland Jens R, Werge T, Daly Mark J, Sullivan Patrick F, O'Donovan Michael C, Kraft P, Hunter DJ, Adank M, Ahsan H, Aittomäki K, Baglietto L, Berndt S, Blomquist C, Canzian F, Chang-Claude J, Chanock SJ, Crisponi L, Czene K, Dahmen N, Silva IdS, Easton D, Eliassen AH, Figueroa J, Fletcher O, Garcia-Closas M, Gaudet MM, Gibson L, Haiman CA, Hall P, Hazra A, Hein R, Henderson BE, Hofman A, Hopper JL, Irwanto A, Johansson M, Kaaks R, Kibriya MG, Lichtner P, Lindström S, Liu J, Lund E, Makalic E, Meindl A, Meijers-Heijboer H, Müller-Myhsok B, Muranen TA, Nevanlinna H, Peeters PH, Peto J, Prentice RL, Rahman N, Sánchez MJ, Schmidt DF, Schmutzler RK, Southey MC, Tamimi R, Travis R, Turnbull C, Uitterlinden AG, van der Luijt RB, Waisfisz Q, Wang Z, Whittemore AS, Yang R, Zheng W, Kathiresan S, Pato M, Pato C, Tamimi R, Stahl E, Zaitlen N, Pasaniuc B, Belbin G, Kenny EE, Schierup MH, De Jager P, Patsopoulos NA, McCarroll S, Daly M, Purcell S, Chasman D, Neale B, Goddard M, Visscher PM, Kraft P, Patterson N, Price AL. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. The American Journal of Human Genetics. 2015;97(4):576-92.

29. Mars NJ, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, Ahola-Olli A, Kurki M, Karjalainen J, Palta P, Neale BM, Daly M, Salomaa V, Palotie A, Widén E, Ripatti S. Polygenic and clinical risk scores and their impact on age at onset of cardiometabolic diseases and common cancers. 2019:727057.

30. Richardson TG, Harrison S, Hemani G, Davey Smith G. An atlas of polygenic risk score associations to highlight putative causal relationships across the human phenome. eLife. 2019;8:e43657.

31. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. Nat Genet. 2018;50(12):1728-34.

32. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32(5):377-89.

33. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. Stat Med. 2017;36(29):4705-18.

34. Snyder PJ, Kopperdahl DL, Stephens-Shields AJ, Ellenberg SS, Cauley JA, Ensrud KE, Lewis CE, Barrett-Connor E, Schwartz AV, Lee DC, Bhasin S, Cunningham GR, Gill TM, Matsumoto AM, Swerdloff RS, Basaria S,

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Diem SJ, Wang C, Hou X, Cifelli D, Dougar D, Zeldow B, Bauer DC, Keaveny TM. Effect of Testosterone Treatment on Volumetric Bone Density and Strength in Older Men With Low Testosterone: A Controlled Clinical Trial. JAMA Intern Med. 2017;177(4):471-9.

35. Traylor M, Rutten-Jacobs LCA, Holliday EG, Malik R, Sudlow C, Rothwell PM, Maguire JM, Koblar SA, Bevan S, Boncoraglio G, Dichgans M, Levi C, Lewis CM, Markus HS. Differences in Common Genetic Predisposition to Ischemic Stroke by Age and Sex. 2015;46(11):3042-7.

36. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, Duncan L, Perry JRB, Patterson N, Robinson EB, Daly MJ, Price AL, Neale BM, ReproGen C, Psychiatric Genomics C, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control C. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236-41.

37. van Rheenen W, Peyrot WJ, Schork AJ, Lee SH, Wray NR. Genetic correlations of polygenic disease traits: from theory to practice. Nature reviews Genetics. 2019;20(10):567-81.

38. Choi MH, Yoo YS, Chung BC. Biochemical Roles of Testosterone and Epitestosterone to 5α-Reductase as Indicators of Male-Pattern Baldness. J Invest Dermatol. 2001;116(1):57-61.

39. Gann PH, Hennekens CH, Ma J, Longcope C, Stampfer MJ. Prospective Study of Sex Hormone Levels and Risk of Prostate Cancer. JNCI: Journal of the National Cancer Institute. 1996;88(16):1118-26.

40. Brand JS, van der Schouw YT, Dowsett M, Folkerd E, Luben RN, Wareham NJ, Khaw KT. Testosterone, SHBG and differential white blood cell count in middle-aged and older men. Maturitas. 2012;71(3):274-8.

41. Palmisano BT, Zhu L, Eckel RH, Stafford JM. Sex differences in lipid and lipoprotein metabolism. Molecular Metabolism. 2018;15:45-55.

42. McHenry J, Carrier N, Hull E, Kabbaj M. Sex differences in anxiety and depression: Role of testosterone. Front Neuroendocrinol. 2014;35(1):42-57.

43. O'Connor DB, Archer J, Wu FCW. Effects of Testosterone on Mood, Aggression, and Sexual Behavior in Young Men: A Double-Blind, Placebo-Controlled, Cross-Over Study. The Journal of Clinical Endocrinology & Metabolism. 2004;89(6):2837-45.

44. Roy CN, Snyder PJ, Stephens-Shields AJ, Artz AS, Bhasin S, Cohen HJ, Farrar JT, Gill TM, Zeldow B, Cella D, Barrett-Connor E, Cauley JA, Crandall JP, Cunningham GR, Ensrud KE, Lewis CE, Matsumoto AM, Molitch ME, Pahor M, Swerdloff RS, Cifelli D, Hou X, Resnick SM, Walston JD, Anton S, Basaria S, Diem SJ, Wang C, Schrier SL, Ellenberg SS. Association of Testosterone Levels With Anemia in Older Men: A Controlled Clinical Trial. JAMA Intern Med. 2017;177(4):480-90.

45. Connallon T, Matthews G. Cross-sex genetic correlations for fitness and fitness components: Connecting theoretical predictions to empirical patterns. Evol Lett. 2019;3(3):254-62.

46. Vihma V, Naukkarinen J, Turpeinen U, Hämäläinen E, Kaprio J, Rissanen A, Heinonen S, Hakkarainen A, Lundbom J, Lundbom N, Mikkola TS, Tikkanen MJ, Pietiläinen KH. Metabolism of sex steroids is influenced by acquired adiposity—A study of young adult male monozygotic twin pairs. The Journal of Steroid Biochemistry and Molecular Biology. 2017;172:98-105.

47. Eriksson J, Haring R, Grarup N, Vandenput L, Wallaschofski H, Lorentzen E, Hansen T, Mellström D, Pedersen O, Nauck M, Lorentzon M, Nystrup Husemoen LL, Völzke H, Karlsson M, Baumeister SE, Linneberg A, Ohlsson C. Causal relationship between obesity and serum testosterone status in men: A bi-directional mendelian randomization analysis. PLoS One. 2017;12(4):e0176277.

48. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-R98.

49. Suderman M, Simpkin A, Sharp G, Gaunt T, Lyttleton O, McArdle W, Ring S, Davey Smith G, Relton C. Sex-associated autosomal DNA methylation differences are wide-spread and stable throughout childhood. 2017:118265.

50. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-9.

51. Vermeulen A, Verdonck L, Kaufman JM. A Critical Evaluation of Simple Methods for the Estimation of Free Testosterone in Serum. The Journal of Clinical Endocrinology & Metabolism. 1999;84(10):3666-72.

52. Emadi-Konjin P, Bain J, Bromberg IL. Evaluation of an algorithm for calculation of serum "Bioavailable" testosterone (BAT). Clin Biochem. 2003;36(8):591-6.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.23.21255981; this version posted April 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

53. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars NJ, Aguirre M, Venkataraman GR, Wainberg M, Ollila HM, Pirruccello JP, Qian J, Shcherbina A, Rodriguez F, Assimes TL, Agarwala V, Tibshirani R, Hastie T, Ripatti S, Pritchard JK, Daly MJ, Rivas MA. Genetics of 38 blood and urine biomarkers in the UK Biobank. 2019:660506.

54. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nature Communications. 2017;8(1):1826.

55. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2015;44(D1):D457-D62.

56. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48(7):709-17.

57. Benner C, Havulinna AS, Järvelin MR, Salomaa V, Ripatti S, Pirinen M. Prospects of Fine-Mapping Trait-Associated Genomic Regions by Using Summary Statistics from Genome-wide Association Studies. Am J Hum Genet. 2017;101(4):539-51.

58. Leinonen JT, Chen Y-C, Pennonen J, Lehtonen L, Junna N, Tukiainen T, Panula P, Widén E. LIN28B affects gene expression at the hypothalamic-pituitary axis and serum testosterone levels. Sci Rep. 2019;9(1):18060.

59. Sivakumaran S, Agakov F, Theodoratou E, Prendergast James G, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson James F, Campbell H. Abundant Pleiotropy in Human Complex Diseases and Traits. The American Journal of Human Genetics. 2011;89(5):607-18.

60. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, Marchini J, Martens JHA, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N, Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang S-Y, Wheeler E, Wilder SP, Iotchkova V, Moore C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D, Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martín D, Watt S, Yang Y, Guigo R, Beck S, Paul DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS, Soranzo N. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell. 2016;167(5):1415-29.e19.

61. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. The MR-Base platform supports systematic causal inference across the human phenome. eLife. 2018;7:e34408.

62. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017;46(6):1734-9.