
Publisher's PDF, also known as Version of record
License (if available): CC BY
Link to published version (if available): 10.1016/j.socscimed.2018.11.004

Link to publication record in Explore Bristol Research
PDF-document

This is the final published version of the article (version of record). It first appeared online via Elsevier at https://www.sciencedirect.com/science/article/pii/S0277953618306385. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights
This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore-bristol-research/ebr-terms/
Testosterone, risk, and socioeconomic position in British men: Exploring causal directionality

Amanda Hughes\textsuperscript{a,b,*}, Meena Kumari\textsuperscript{b}

\textsuperscript{a} MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, United Kingdom
\textsuperscript{b} Institute for Social and Economic Research, University of Essex, United Kingdom

\begin{abstract}
Lower testosterone levels in men are observationally associated with worse health, but it is unclear whether they contribute to well-established social gradients in health. Mendelian Randomization studies suggest positive testosterone-health associations may not be causal, with some intervention studies suggesting testosterone administration could be harmful. Since testosterone is rarely measured in general population studies, very little is known about how testosterone varies by social position. Differences by education and household income in British men aged 60-64y were recently reported, but it is unclear whether this reflects an influence of socioeconomic position (SEP) on testosterone, influence of testosterone on SEP, or confounding.

In the UK Household Longitudinal Study, a nationally-representative survey of UK adults, we examine social differences in testosterone in 3663 men aged 16-97y in 2010-12. We consider diverse dimensions of SEP: education, employment status, equivalized household income and personal earnings. Multivariable regression is used to explore social differences in testosterone across the adult life-span (16-97y). Secondly, Mendelian Randomization (MR), an approach which uses gene variants as instrumental variables for endogenous exposures, is used to investigate causal directionality. We examine associations with risk-taking, a plausible mediator of testosterone-SEP associations.

In observational models no social differences in testosterone are seen, but MR models suggest a positive influence of testosterone on earnings (increase in log-transformed monthly earnings (GBP) per standard deviation increase in testosterone: 0.51, 95\%CI: 0.03,1.05, \(p = 0.07\)) and probability of being in work (probit coefficient:0.25, 95\%CI: 0.01,0.51, \(p = 0.06\)). Though MR estimates are less precise, results are consistent with previous literature linking testosterone with labour market success. The discrepancy may reflect suppression of observational associations by factors positively correlated with testosterone and negatively correlated with SEP, or indicate an influence of typical lifetime testosterone, which may be better indexed by genetic variants than by single testosterone measurements subject to noise.

\end{abstract}

\section{Introduction}

Lower circulating testosterone in men is consistently related to worse health in observational studies, including cardiovascular disease (Kloner et al., 2016), mortality (Khaw et al., 2007), and Alzheimer's disease (Lv et al., 2016). However, important questions remain regarding the causal mechanisms linking testosterone and health. Studies utilizing genetic variation in testosterone suggest high testosterone does not causally benefit health (Eriksson et al., 2017; Haring et al., 2013; Zhao et al., 2014, 2016a; Svarberg et al., 2014; Schooling et al., 2018). Although evidence is mixed (Cheetham et al., 2017), some studies of testosterone supplementation have suggested testosterone administration could increase risk of cardiovascular events (Albert and Morley, 2016), leading the FDA to require warnings labels on supplements (FDA, 2014). With a few exceptions (Mazur, 2009; Svarberg et al., 2003; Harman et al., 2001), descriptions of age differences come from age-restricted populations (Yeap et al., 2007; Lapauw et al., 2008; Gapstur et al., 2007; Mohr et al., 2005; Liu et al., 2007; Orwoll et al., 2006), meaning social patterning across the life-course is yet to be investigated. Meanwhile, very little is known about social inequalities in testosterone, which is rarely measured in general population surveys. A recent analysis in the National Survey of Health and Development (NSHD) found social differences in men's circulating testosterone at age 60-64y, with lower testosterone for men with lower income and fewer

\begin{itemize}
\item Lower testosterone levels in men are observationally associated with worse health.
\item Mendelian Randomization studies suggest positive testosterone-health associations may not be causal.
\item Social differences in testosterone were examined in the UK Household Longitudinal Study.
\item Multivariable regression was used to explore social differences in testosterone across the adult life-span.
\item Mendelian Randomization (MR) was used to investigate causal directionality.
\item In observational models no social differences in testosterone were seen.
\item MR models suggested a positive influence of testosterone on earnings.
\item The discrepancy may reflect suppression of observational associations by factors.
\end{itemize}
educational qualifications (Bann et al., 2015). This raises the possibility that, if testosterone-health associations are partly causal, testosterone could reflect an overlooked mechanism contributing to social inequalities in health.

Testosterone and socioeconomic position (SEP) could be associated in a number of directions and ways. Thus, testosterone could be causally influenced by aspects of SEP, but social patterning of testosterone could also reflect impact of testosterone on SEP. Finally, testosterone-SEP relationships could be confounded, including by health. Using data on men aged 16-98y from a large, nationally-representative British survey, we examine social patterning of testosterone across the adult life-span, with age-curves described for groups of income and education. Multivariable regression is used to examine age-adjusted associations with income, employment status, education, and self-assessed risk-taking behaviour, proposed as a mediator of testosterone-SEP associations. As this analysis was conducted without a priori assumptions about the principal causal direction of associations, two income measures are used. If socioeconomic conditions primarily influence testosterone via the mechanisms discussed below, associations should be strongest with equivalized net household income, but if testosterone primarily influences SEP via the mechanisms discussed below, associations should be strongest with gross earnings. Finally, Mendelian Randomization (Burgess and Thompson, 2015) is used to investigate causal influence of testosterone on SEP, using gene variants (rs12150660, rs6258, rs934505) as exogenous genetic instruments for circulating testosterone.

1.1. Impact of testosterone on socioeconomic position

For human and non-human primates, testosterone is thought to play a role in advancing and maintaining status by encouraging ‘dominance behaviour’, which aims to enhance one’s status compared to competitors (Mazur, 1985; Archer, 2006). While early work focused on aggression in dominance behaviour (Mazur and Booth, 1998), recent work suggests that in humans, testosterone plays a more nuanced role in status promotion, by encouraging either aggressive or prosocial behaviour depending on the context (Dreher et al., 2016; Carre and Archer, 2018). Moreover, researchers now recognise human aggression as something which can take ‘purely psychological or even economic forms, rather than being overtly violent’ (Eisenegger et al., 2011). Supporting the idea that testosterone is conducive to ‘economic aggression’, experimental work has found positive associations between testosterone and financial risk taking (Cuena et al., 2015; Nofsinger et al., 2018; Mehta et al., 2015), although null associations are also reported (Aicella et al., 2015). There is evidence that these behavioural implications could extend beyond the laboratory, potentially with relevance to longer-term socioeconomic position. A study of male executives found higher testosterone was associated with having more subordinates (Sherman et al., 2016), while other studies find that testosterone in men is associated with self-employment, a ‘riskier’ strategy than standard employment (Greene et al., 2014; Nicolau et al., 2018), although null associations with self-employment have also been reported (van der Loos et al., 2013). Studies of male financial traders report that daily profits were predicted by morning testosterone (Coates and Herbert, 2008) and 2D:4D ratio, believed to reflect prenatal testosterone exposure (Coates et al., 2009), with authors explaining these positive associations of testosterone and profits as a function of greater risk tolerance (Coates and Gurnell, 2017). If riskier behaviour can lead to better financial outcomes, this raises the possibility of cumulative influence on long-term social position via wealth (Stanton, 2017). However, behavioural attributes which make one a successful financial trader may not be beneficial in other professions, and whether testosterone is more widely conducive to financial success is not clear. Two recent papers examined this by looking at plausible indicators of in utero testosterone exposure (2D:4D ratio, or sex of a twin) in relation to earnings in adulthood. One found that lower 2D:4D ratio, thought to correspond to high in utero testosterone, predicted greater wages for men and women. However, there was some evidence of nonlinear effects (Nye et al., 2017). The other found that male sex of the twin (corresponding to greater prenatal testosterone exposure) predicted higher earnings for men but lower earnings for women (Gienlen et al., 2016). However, the consequences of in utero and adult circulating testosterone may differ (Hönekopp et al., 2007), meaning investigation of relationships with circulating testosterone is warranted. Given involvement of anabolic hormones in growth and evidence linking taller stature to SEP in men (Tyrrell et al., 2016), an additional pathway linking testosterone to SEP may operate through height.

1.2. Impact of socioeconomic position on testosterone

Social patterning of testosterone could also be explained by a causal impact of SEP on testosterone, through several mechanisms. Firstly, human and animal studies have shown that stress, especially chronic stress, can lower circulating testosterone by influencing both production and secretion (Chichinadze and Chichinadze, 2008). An impact on testosterone of SEP might be therefore expected, given evidence that psychosocial stress associated with socioeconomic adversity can impact health via other biological pathways (Seeman et al., 2010). Secondly, analogously to other blood biomarkers (Davillas et al., 2017), an impact of SEP on testosterone could operate via socially-patterned health behaviours. Most plausible is adiposity, in high-income countries positively associated with socioeconomic disadvantage, and which lowers circulating testosterone (Gapstur et al., 2007; Cooper et al., 2015). Conversely, smoking is strongly linked to disadvantage but raises testosterone (Mohr et al., 2005; Zhao et al., 2016b), so could work against these effects or even produce a ‘reverse gradient’. Meanwhile, experimental research shows circulating testosterone is sensitive to the social environment, in particular where an individual’s status is threatened: testosterone changes in response to competition depending on outcome, rising in the winner compared to the loser (Archer, 2006; Mazur and Booth, 1998; Geniole et al., 2017). Since testosterone also seems to predict success in some competitive situations, a feedback loop or ‘winner effect’ has been hypothesised whereby experience of past success increases probability of future success (Coates and Gurnell, 2017). There is evidence for such an effect among humans in experimental settings and in sports matches (Page and Coates, 2017). However, as with impact on financial risk taking, it remains to be shown that these processes occur widely enough to plausibly contribute to population-level SEP differences in testosterone.

Several factors could give rise to confounded associations of SEP with testosterone. Lower testosterone is consistently linked to poor health, but recent evidence suggests it may reflect poor health rather than driving it (Eriksson et al., 2017; Haring et al., 2013; Zhao et al., 2014, 2016a; Svrberg et al., 2014). Testosterone-SEP association could therefore reflect more general associations of disadvantaged socioeconomic position and poor health (Nandi et al., 2014; Stringhini et al., 2010) without contributing to them. Another possible confounder is household composition, especially where household income-based measures are used to index SEP. Testosterone is lower in partnered than single men (Booth and Dabb, 1993; Gettler et al., 2013), and fathers than childless men (Pollet et al., 2013; Gettler et al., 2011), although contradictory results have been reported (Mazur, 2014). These findings are explained using evolutionary theory: with pair-bonding, an individual male’s priorities change from competing against other males for mates to cooperating with the existing one, such that a hormonally-mediated shift towards cooperative behaviour is adaptive (Mazur, 2014). However, since both cohabitating partnership status and children are important determinants of an adult’s equalised household income, these factors may bias associations of testosterone and frequently used income-based SEP measures. For this reason, our observational models of household income adjust for these factors. A further complication concerns the recently developed dual hormone
hypothesis (Mehta and Prasad, 2015). This proposes that in humans, testosterone's influence on behaviour depends on cortisol, such that testosterone only increases status-promoting behaviours when cortisol is low (Carre and Archer, 2018; Mehta et al., 2015). Supporting this, many of the recent studies supporting an influence of testosterone on risk behaviour (Cueva et al., 2015; Nofsinger et al., 2018; Mehta et al., 2015) and status position (Sherman et al., 2016) find this only applies for low-cortisol individuals. However, diurnal variation in cortisol itself differs by SEP (Kumari et al., 2010; Karlamangla et al., 2013), again raising the possibility of confounding. More work will be required to explore mechanisms linking cortisol, testosterone, and SEP.

2. Methods

2.1. Participants

2.1.1. Study participants

The UKHLS is an annual longitudinal survey of over UK 40,000 households. It consists of a larger General Population Sample (GPS), a stratified clustered random sample of households representative of the UK population which joined in 2009–10, and a smaller component from the pre-existing British Household Panel Survey (BHPS) (Knies, 2015). Blood samples taken during a nurse visit approximately five months after the main wave 2 interview (GPS participants) or wave 3 interview (BHPS participants) in England, Wales and Scotland (not Northern Ireland). In the second year of data collection for the GPS sample, eligibility was restricted to 0.81 of primary sampling units (PSUs) in England (McFall et al., 2014). Eligibility criteria required that these participants were aged 16+, had participated fully in the previous main interview in English, and did not have HIV or a clotting disorder (McFall et al., 2014). Approximately 50% of individuals meeting these criteria were successfully contacted (N = 20644). Of these, blood samples were obtained from 13,238, largely due to non-consent for a blood sample. For genotyping, participants needed to give additional consent and be of white ethnicity; 9944 met these criteria. Following standard genetic quality control procedures (see below) for pairs of individuals related by r² = 0.20 or more one was randomly chosen for inclusion. Of the 4358 men remaining, exclusion for missing testosterone (N = 48), 0 values for inverse-probability weights, 2 participants taking anabolic steroids and one outlier for the polygenic score left a final sample size of 3663 men.

2.2. Measures

2.2.1. Testosterone

Total serum testosterone was measured by an electrochemiluminescent immunoassay on the Roche Modular E170 analyser; QC checks showed the intra and inter assay coefficient of variation were less than 4% (Benzeval et al., 2014). Observations below the detection limit of 1 nmol/L were recoded to 0.5 nmol/L, and those above the maximum detectable value of 48 nmol/L to 48 nmol/L. Testosterone measurements approximated a normal distribution, so were left un-transformed. Since almost all women were below the detection limit of 1 nmol/L, analysis was restricted to men.

2.2.2. Genotyping and genetic variants

Samples were genotyped using the Illumina HumanCore Exome and imputation carried out in Minimac 5-12-29 to the European component of 1000 genomes (Prins et al., 2017). Individuals were dropped if genetic information was discordant with stated ethnicity, sex, or relatedness to other sample members. For twins, the twin with the lower call rate was excluded. Pre-imputation quality control removed SNPs with a minor allele frequency of <1%, call rate threshold <98%, Hardy-Weinberg Equilibrium \( p < 10^{-4} \) or cluster separation score < 0.4. Three genetic variants were used as instruments for circulating testosterone: rs12150660 and rs6258 in the SHGB gene on chromosome 17, and rs5934505 near FAM9B on the X chromosome. These were identified in 2011 genome-wide analysis by Ohlsson et al. to independently explain 2.3%, 0.9% and 0.6% respectively of the variance of serum testosterone (Ohlsson et al., 2011) and have been used in previous MR analyses of testosterone in relation to health\( (4, 5) \). An externally-weighted polygenic score (PGS) for testosterone was calculated using beta values for per-minor allele association with testosterone from Ohlsson’s GWAS and used as a single instrumental variable in analyses. In UKHLS, rs6258 and rs5934505 were genotyped, while rs12150660 was imputed (imputation \( r^2 = 0.94 \) ). A subsequent, smaller GWAS in a sample of older men at increased risk of prostate cancer (Jin et al., 2012) reported slightly different results. We use betas from Ohlsson’s GWAS to construct the polygenic score because of its larger size, and because the nature of the sample (a collection of general population surveys) is more appropriate to the current study population of UKHLS.

There was one clear outlier for the PGS, the sole homozygote for the testosterone-lowering allele of rs6258. Their PGS was more than 7 standard deviations below the sample mean (with the next lowest value around 4 SD below the mean), who also had very low testosterone (around the 14th centile for the five-year age band centered at his age). To ensure robustness of results, he was excluded from the main analyses, but his contribution examined in a robustness check.

2.2.3. Socioeconomic position measures

Two measures of monthly income were considered, based on self-report information from the main interview closest to the biomedical assessment: equalized net household income and gross earnings. To deal with likely errors, the top and bottom 0.5% of observations were removed. For age-couples, variables were divided into tertiles within 5-year age-bands. For further analysis, variables were log-transformed (after adding 1 to 0-values). Employment status was self-reported at the main interview corresponding to the nurse visit. From this, binary indicators were constructed: whether in work (employed or self-employed), whether self-employed compared to anything else, and whether self-employed conditional on working. Educational qualifications, from questionnaire information, were categorised as no qualifications, qualifications below degree, or university degree or equivalent. We followed the standardization procedure of Fiorito et al. to account for cohort differences in education (Fiorito et al., 2017). Education was first categorised into the three groups above, then standardized by 5-year age band to produce a continuous score between 0 and 1, with 1 corresponding to the most education relative to peers. This was to reflect the generationally changing ‘meaning’ of university education (Galobardes et al., 2006). For younger participants more likely to have attended university, a degree may not capture the same background characteristics, nor afford the advantages, as several decades ago.

2.2.4. Self-reported risk tolerance

Participants rated their willingness to take risks at wave 18 of BHPS (BHPS participants) or wave 1 of UKHLS (GPS participants). The wording of the questions was identical, ‘are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?’ However, the wording and numerical range of possible answers differed: BHPS participants answered from 0 ‘unwilling to take risks’ to 10 ‘fully prepared to take risks’, but GPS participants from 1 ‘avoid taking risks’ to 10 ‘fully prepared to take risks’. While the distribution of answers was approximately normal for both subsamples, the shorter numerical range offered to BHPS participants resulted in a significantly higher mean value (5.96 vs 5.60, adjusting for age and age²).

2.2.5. Confounders

In an observational analysis, a confounder is a factor which causally influences both exposure and outcome. In instrumental variable analysis (including MR) a confounder is a factor which causally influences the instrument (and therefore the predicted value of the exposure based
on the instrument) and the outcome. Factors influencing the outcome and the part of the exposure not determined by the instrument do not bias the IV estimate, which is the key strength of instrumental variable analysis. Meanwhile, adjusting for variables which are not confounders but may be colliders (that is, jointly determined by exposure and outcome) can introduce bias, a phenomenon receiving increasing attention (Davies et al., 2018). To balance these considerations, plausible confounders of observational but not IV estimates (smoking, adiposity, timing of testosterone measurement, region, medications) were included in observational but not IV models. IV estimates adjusted for observational confounders are presented in a supplementary table for reference only.

Age was from self-report at the biomedical assessment; a squared term was included to allow for nonlinearity in the association of testosterone and age. Time of day was included as a continuous measure of the time when the assessment started (24hr format, range 7–21). Adiposity was indexed by percent body fat, measured by a nurse using Tanita digital floor scales (McFall et al., 2014). This is increasingly considered a better adiposity measure than BMI, which not distinguish body composition (Davillas and Bezenal, 2016). Height was measured by a nurse with a portable stadiometer(64). Participants reported prescribed medications at the nurse visit, checked by the nurse from medication containers. Use of beta-blockers, statins, non-steroidal anti-inflammatory medications, and antidepressants was classified using recorded BNF codes. Smoking status was classified as never smoker, ex-smoker, current ≤ 10/day, current 11–20/day, current > 20/day, using self-reported information from wave 2 for all participants. At each wave, participants rated their overall health as excellent/very good/fair/poor. Information was taken from the annual visit closest to the biomedical assessment and analysed as continuous. In genetic analyses the first 10 genetic ancestry principal components were included to account for possible population stratification. UK Government office region (GOR) was identified from participant postcodes and classified as North East/North West/Yorkshire and Humberside/East Midlands/West Midlands/East Anglia/London/South East/South West/Scotland/Wales. Because of the difference in measurement of self-assessed risk taking between BHPS and GPS participants, and because income measurements for BHPS participants were taken a year later, a variable was included indicating GPS/BHPS subsample. Participants were assigned 1 rather than 0 for partnership status if they reported being married, in a civil partnership, or cohabiting with a partner. At all waves participants were asked about dependent children aged under 18y, from which a binary indicator was derived.

2.3. Analysis

Social differences in testosterone across the life-course were explored by calculating age-curves of testosterone separately by equivalentized household income tertiles and educational qualifications. Since income is strongly influenced by age, tertiles were defined separately by 5-year age bands. Testosterone was regressed on 5-year age band, income tertile or educational qualifications, and time of day of blood sampling, with an interaction of age-band and the SEP measure. Predicted values were extracted for each combination of age-band and SEP, with time of day set to the mean of 2.40pm. Multivariable regression models used linear regression for continuous dependent variables (equivalized net household income, gross earnings, risk-taking, age-standardized education) and probit regression for binary variables (employment, self-employment, partnership and children). For Mendelian Randomization, instrumental variable analysis was performed using ivregress for continuous outcomes and ivprobit for binary outcomes. Models of equivalized household income were re-run adjusting for partnership and dependent children. All OLS models adjusted for age, age2, % body fat, smoking status, self-rated health, GPS/BHPS subsample, beta-blockers, statins, NSAIDs, antidepressants, government office region and first ten principal components, while MR models adjusted for age, age2, and the first ten principal components only. All models used the maximum sample for the particular regression, meaning sample size differed slightly between models. To account for non-response and sampling bias, all models were inverse-probability weighted using blood weights from the nurse visit (UKHLS), and STATA’s svyset command used to account for clustering and sample stratification. Coefficients are presented per standard deviation increase in testosterone (6.0 nmol/L).

2.3.1. Robustness checks

All analyses were repeated with the age range restricted to 22–70, since there is evidence that total testosterone correlates less well with bioavailable testosterone in men older than 70(16), and since below 22 low earnings may reflect higher education participation. For MR analyses, standard robustness checks for weak instrument bias were performed. Models were run to check whether the PGS predicted other relevant factors, namely smoking, percent body fat, height, medications use, self-rated health and General Health Questionnaire (GHQ) score, which measures psychological distress. Models were repeated including the PGS outlier and results compared.

3. Results

3.1. Sample description

Descriptive characteristics of the sample, overall and by testosterone tertiles, are shown in Table 1. Adjusted for age, age2 and the first 10 principal components, men with an additional minor allele for rs12150660 (T) had testosterone 1.26 nmol/L higher (p < 0.001), men with an additional minor allele (T) for rs6258 had testosterone 2.27 nmol/L lower (p = 0.006), and men with an additional minor allele (C) for rs5934505 had testosterone 1.42 nmol/L higher (p < 0.001). The r2 from a regression of testosterone on the PGS only showed the PGS explained 2.6% of variance in testosterone. Regressing testosterone on each SNP in turn showed that rs12150660, rs6258 and rs5934505 respectively explained 1.7%, 0.2% and 1.1% of the variance in testosterone. Adjusted for age and age2, time of day of sampling was inversely and significantly related to testosterone (−0.38 nmol/L/hour, p < 0.001); a model using dummy variables found no evidence for nonlinearity in effects. Adjusted for age, age2, and each other, testosterone was significantly negatively related both to presence of dependent children (−0.91 nmol/L, p = 0.001) and cohabiting with a partner (−0.82 nmol/L, p = 0.002).

3.2. Age curves of testosterone by SEP group

Descriptive age curves were calculated using STATA’s margins command, with all results presented for the mean time of day (2.40pm). For all men (Fig. 1) there was clear age-related decline as expected, from 18.8 (95%CI:17.4–20.1) at 16–20 to 12.6(10.6–14.5) at 86–97. Age curves of testosterone by SEP groups (age-specific tertiles of equivalized household income (Fig. 2), and educational qualifications (Fig. 3) did not support substantial social differences in testosterone. In both cases, confidence intervals largely overlapped, and differences showed no consistent direction. Taking account of health behaviors by calculating marginal SEP effects at the mean value of percent body fat and smoking status (Figs. 4 and 5), or of partnership and children (Figs. 6 and 7) did not qualitatively change results.

3.3. OLS and IV models: associations of testosterone with income, education and risk taking

OLS models adjusted for age and age squared, percent body fat, smoking status, government office region and GPS/BHPS sample found no association of testosterone with self-assessed risk-taking behaviour among men aged 16–97y (Table 2). There were no associations with the
age-standardized measure of education, and no association with probability of working or of self-employment (Table 2). Coefficients for all income measures were negative, non-significant, and close to zero. Repeating analysis in the age-restricted sample (Table 3) showed similar results. Since heterogeneity in behaviourally-mediated impact of testosterone on earnings was plausible, quantile regression was used to check for higher testosterone at both distribution tails masking a difference at the mean. This was not supported: coefficients for

Table 1
Descriptive characteristics of analytic sample: Men in UKHLS, 2010–12, N = 3663.

<table>
<thead>
<tr>
<th>Category</th>
<th>All</th>
<th>Tertile 1</th>
<th>Tertile 2</th>
<th>Tertile 3</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Age</td>
<td>54.9(16.1)</td>
<td>55.2(15.4)</td>
<td>55.0(15.7)</td>
<td>54.4(17.2)</td>
<td>3663</td>
</tr>
<tr>
<td>Testosterone</td>
<td>15.4(6.0)</td>
<td>9.4(2.8)</td>
<td>15.1(1.3)</td>
<td>21.9(4.2)</td>
<td>3663</td>
</tr>
<tr>
<td>Percent body fat (%BF)</td>
<td>24.8(8.9)</td>
<td>26.1(8.8)</td>
<td>24.0(8.4)</td>
<td>21.8(6.6)</td>
<td>3349</td>
</tr>
<tr>
<td>Body mass index (BMI)</td>
<td>28.2(4.6)</td>
<td>30.1(5.0)</td>
<td>28.2(4.2)</td>
<td>26.5(3.8)</td>
<td>3550</td>
</tr>
<tr>
<td>Self-assessed risk-taking</td>
<td>5.6(2.4)</td>
<td>5.5(2.4)</td>
<td>5.6(2.5)</td>
<td>5.7(2.4)</td>
<td>3174</td>
</tr>
<tr>
<td>Equivalised household income (monthly £, net)</td>
<td>1720(1034)</td>
<td>1720(979)</td>
<td>1727(1069)</td>
<td>1713(1054)</td>
<td>3626</td>
</tr>
<tr>
<td>Earnings (monthly £, gross)</td>
<td>1377(1610)</td>
<td>1456(1588)</td>
<td>1399(1634)</td>
<td>1274(1576)</td>
<td>3609</td>
</tr>
</tbody>
</table>

Fig. 1. Age variation in total testosterone for men in UKHLS (N = 3663). Values given for the mean time of day of blood sampling (2.40pm).
testosterone at the 5th, 10th, 25th, 50th, 75th, 90th and 95th centile respectively were 0.00(0.00,0.00), 0.00(0.00,0.00), $-0.07(-0.19,0.04)$, 0.00(-0.05,0.06), 0.05 ($-0.00,0.11$), 0.03(-0.01,0.06) and 0.04($-0.00,0.07$). In an attempt to replicate results from the recent analysis (22) reporting descriptive differences by household income and education in total and free testosterone at age 60-64y (unadjusted for health behaviours, medications or household composition), observational models in were repeated in men aged 60-64y with adjustment only for age and time of day (N = 408 to 414). No social differences in testosterone were seen in this subsample ($-0.04$ (95%CI: 0.10,0.02, p = 0.17) for equivalized household income, 0.19(95%CI: 0.15,0.53, p = 0.27) for earnings, and $-0.00$ 95%CI(-0.03-0.02, p = 0.88) for age-standardized education).

IV analyses using the PGS (Table 2) also did not find an association of testosterone with equivalized net household income, with or without adjustment for partnership and children, with age-standardized education, or with odds of self-employment. However, results did suggest association with log-transformed gross earnings (0.51, 95%CI: 0.03-1.05, p = 0.07), and probability of being in work (0.25, 95%CI: 0.01-0.51, p = 0.06). Repeating analysis in the age-restricted sample (Table 3) showed similar results. Including the PGS outlier strengthened associations with log-transformed gross earnings (0.65, CI: 0.05–1.26, p = 0.03) and likelihood of being in work (0.31, CI: 0.05–0.57, p = 0.02), which again were similar in the age-restricted model (Tables S2 and S3). These IV associations were also stronger (for gross earnings, 0.62, CI: 0.10–1.14, p = 0.02, for likelihood of working 0.32, CI:0.05–0.59, p = 0.02) when adjusted for observational confounders (Table S4).

Adjusting for age, age$^2$ and the first ten principal components, standard robustness checks for instrumental variable analysis confirmed the PGS was not a weak instrument (minimum first-stage F-statistic of 88.5 for risk-taking, which had the smallest sample). Checks for whether the PGS predicted other relevant factors (Table 4) showed it did not significantly predict percent body fat, self-rated health,
psychological distress, height, or smoking. Of medications, it did predict lower probability of beta-blocker use (OR: 0.87, p = 0.01).

To test whether residual confounding in smoking could have caused suppression of the association of testosterone and earnings in the OLS models, observational models for earnings were repeated using a cruder smoking measure of never/ex-/current. Results were very similar (−0.06, 95%CI: -0.16-0.04 in the full sample, 0.00, 95%CI: 0.10-0.10 in the age-restricted sample), suggesting finer-grained inaccuracies in smoking were not greatly influencing estimates. To test the ability of the instrument to correct for confounding, OLS and IV models were rerun with minimal adjustment, including only age, age² and principal ancestry components (Table S1). This induced negative confounding by smoking and time of day of blood sampling, which have opposite associations with earnings and testosterone. With this specification, a significantly negative association of testosterone and earnings was seen in the OLS model (−0.16, 95%CI: 0.27 to −0.05), in contrast to the positive association in the IV model (0.51, 95%CI: 0.03 - 1.05).

A rough power estimate (mRnd) indicated ample power for the MR analysis despite the modest sample size, reflecting the strength of the instrument. For example, assuming a true causal effect for the testosterone-log earnings association of only 0.2, estimated power was 91%. However, since the tools currently available to estimate power of a 2SLS analysis assume a simple random sample, this is likely to be an overestimate, as the UKHLS is a complex survey with stratification, clustering and for which inverse-probability weights were applied.

4. Discussion

This analysis found no evidence of social differences in total testosterone by household income, education, or probability of self-employment among men aged 16–97. For these aspects of SEP, conclusions did not differ using descriptive age-curves, age-adjusted observational models, or genetic instrumental variable analysis. Results are therefore at odds with descriptive differences by household income and

Fig. 4. Men’s testosterone by age and household income tertile (N = 3286). Values given for mean time of blood sampling, smoking, and percent body fat. Income tertiles calculated within each 5-year age band.

Fig. 5. Men’s testosterone by age and educational qualifications in UKHLS (N = 3241). Values given for mean time of blood sampling, smoking, and percent body fat.
education in total and free testosterone at age 60-64y recently reported in the NSHD (Bann et al., 2015). Since no differences were seen in UKHLS participants aged 60-64y when adjusting only for age and time of day, this suggests sample differences in general social representativeness, or aspects of health likely to affect testosterone, may be responsible. Substantial differences between sample means of testosterone at similar ages (Ruiz and Kumari, 2017) support this interpretation.

For the remaining SEP measures, observational and MR models produced notably different results. MR estimates were less precise owing to the greater power required for an IV analysis. Nevertheless, in contrast to observational models, these suggest a positive association of testosterone with earnings and with probability of working, consistent with a causal influence of testosterone on these aspects of SEP. Both associations increased when an outlier with a very low PGS was included, which may indicate a degree of pleiotropy for rs6258. Alternatively, since his testosterone was very low, this could reflect a nonlinearity whereby the impact of testosterone is greatest at the lower end of the testosterone range, in which case results of the main analysis should be considered conservative. Although MR associations were stronger adjusting for observational confounders, results should be treated with caution as the discrepancy may reflect collider bias. Interestingly, there was little evidence of an association with risk-taking. Although the self-reported measure has clear limitations, this result runs counter to the interpretation of some researchers that risk-taking mediates financial benefit of testosterone. Similarly, the PGS did not predict self-rated health. This suggests the mechanisms linking testosterone and earnings are not explained by overall health, but may also reflect limitations of self-report measures in the study of socioeconomic inequalities in health (Layes et al., 2012).

One explanation for the contrasting results of OLS and MR models is downward bias in observational models due to suppression, or negative confounding, by factors negatively associated with earnings and positively associated with testosterone. One possible source of suppression

Fig. 6. Men’s testosterone by age and household income tertile in UKHLS (N = 3626). Values given for mean time of blood sampling, partnership, and children. Income tertiles calculated separately within each 5-year age band.

Fig. 7. Men’s testosterone by age and educational qualifications in UKHLS (N = 3569). Values given for the mean time of blood sampling, partnership, and children.
is unobserved heterogeneity in smoking. This was measured using self-report information from wave 2, and may therefore have been influenced by social desirability bias or been out of date for some participants. However, observational estimates for earnings were very similar using a cruder smoking measure, suggesting fine-grained inaccuracies in smoking were not responsible for masking associations. Nevertheless, confounding by other factors associated negatively with earnings and positively with testosterone may have affected OLS estimates, whilst being largely corrected for in the MR models. Ability of the PGS to correct for confounders of observational testosterone-SEP associations is supported by comparison of OLS and IV results from the minimally-adjusted models. A second explanation for the contrast between OLS and MR estimates is that an individual's testosterone varies from day to day, but labour market success is plausibly influenced in a cumulative manner by a person's typical circulating testosterone over a lifetime. It should therefore show a clearer relationship with an approximation of an individual's personal mean than with any single testosterone measurement, since the latter will be influenced both by stable factors contributing to inter-individual testosterone differences (e.g. underlying long-term health conditions, and gene variants) and shorter-term sources of variation. The values for genetic associations with testosterone estimated in the GWAS are also subject to noise, since they were estimated using single testosterone measurements in a finite sample. However, since genes do not change within an individual's lifetime, the true associations they approximate relate to the single most stable component of inter-individual variation in testosterone. Thus, the MR estimate represents a (confounded) result of a standard deviation difference in one standard deviation increase in testosterone(23), while the OLS estimate represents a (confounded) result of a standard deviation increase in testosterone on a given day. A clearer association in the MR models is therefore perhaps not surprising. Another way to approximate the stable component of testosterone would be to use an average of several testosterone measurements per individual, and studies with data on earnings and repeated testosterone measurements should seek to do this.

Importantly, while OLS models find no positive evidence for an influence of the SEP measures on testosterone, detectability of such associations may have been influenced by noisiness of the single testosterone measure. Surveys with repeat measurements should seek to investigate this question using averaged values. In any case, results suggest that in British men aged 16-97y, the impact of household income, earnings, education, and employment status on testosterone is not large enough to be detectable with a single testosterone measurement in 3663 men from a representative population survey. Results therefore suggest that impact of socioeconomic position on testosterone

---

**Table 2**

Association of men’s serum testosterone (nmol/L) with social and behavioural factors in UKHLS 2010-12.

<table>
<thead>
<tr>
<th>Coefficients per SD of testosterone (6.0 nmol/L)</th>
<th>OLS Estimates</th>
<th>IV Estimates</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff</td>
<td>95%CI</td>
<td>p</td>
</tr>
<tr>
<td>Results of OLS and ivregress models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equivalised household net&lt;sup&gt;a&lt;/sup&gt;</td>
<td>−0.00</td>
<td>(−0.02 - 0.02)</td>
<td>0.94</td>
</tr>
<tr>
<td>Equivalised household net, adjusted for partnership and children&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−0.00</td>
<td>(−0.02 - 0.01)</td>
<td>0.64</td>
</tr>
<tr>
<td>Earnings, gross&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.06</td>
<td>(−0.17 - 0.04)</td>
<td>0.24</td>
</tr>
<tr>
<td>Earnings, gross, conditional on being in work&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.00</td>
<td>(−0.04 - 0.04)</td>
<td>0.89</td>
</tr>
<tr>
<td>Age-standardized education (range 0–1)</td>
<td>−0.00</td>
<td>(−0.01 - 0.00)</td>
<td>0.34</td>
</tr>
<tr>
<td>Self-assessed risk-taking: range 0–10 (BHPS) or 1–10 (GPS)</td>
<td>0.02</td>
<td>(−0.07 - 0.11)</td>
<td>0.62</td>
</tr>
<tr>
<td>Results of probit and ivprobit models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In work (employed/self-employed)</td>
<td>0.00</td>
<td>(−0.06 - 0.06)</td>
<td>0.96</td>
</tr>
<tr>
<td>Self-employed</td>
<td>0.00</td>
<td>(−0.06 - 0.06)</td>
<td>0.93</td>
</tr>
<tr>
<td>Self-employed, conditional on working</td>
<td>−0.01</td>
<td>(−0.08 - 0.06)</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<sup>a</sup> Testosterone associations with outcomes were estimated separately, adjusting for confounders but not each other.

<sup>b</sup> Adjusted for age, age², time of day of testosterone sampling, % body fat, smoking status, GPS/BHPS subsample, beta-blockers, statins, NSAIDs, antidepressants, government office region and first ten principal components.

<sup>c</sup> Adjusted for age, age² and first ten principal components.

<sup>d</sup> Income measures are log-transformed and based on monthly income.

---

**Table 3**

Association of men’s serum testosterone (nmol/L) with social and behavioural factors in UKHLS 2010–12, restricted age range 22–70.

<table>
<thead>
<tr>
<th>Coefficients per SD of testosterone (6.0 nmol/L)</th>
<th>OLS Estimates</th>
<th>IV Estimates</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff</td>
<td>95%CI</td>
<td>p</td>
</tr>
<tr>
<td>Results of OLS and ivregress models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equivalised household net&lt;sup&gt;e&lt;/sup&gt;</td>
<td>−0.01</td>
<td>(−0.02 - 0.01)</td>
<td>0.66</td>
</tr>
<tr>
<td>Equivalised household net, adjusted for partnership and children&lt;sup&gt;f&lt;/sup&gt;</td>
<td>−0.01</td>
<td>(−0.02 - 0.01)</td>
<td>0.60</td>
</tr>
<tr>
<td>Earnings, gross&lt;sup&gt;e&lt;/sup&gt;</td>
<td>−0.00</td>
<td>(−0.13 - 0.11)</td>
<td>0.87</td>
</tr>
<tr>
<td>Earnings, gross, conditional on being in work&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.01</td>
<td>(−0.04 - 0.05)</td>
<td>0.73</td>
</tr>
<tr>
<td>Age-standardized education (range 0–1)</td>
<td>0.00</td>
<td>(−0.01 - 0.01)</td>
<td>0.99</td>
</tr>
<tr>
<td>Self-assessed risk-taking: range 0–10 (BHPS) or 1–10 (GPS)</td>
<td>0.00</td>
<td>(−0.10 - 0.10)</td>
<td>0.95</td>
</tr>
<tr>
<td>Results of probit and ivprobit models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In work (employed/self-employed)</td>
<td>0.01</td>
<td>(−0.07 - 0.09)</td>
<td>0.72</td>
</tr>
<tr>
<td>Self-employed</td>
<td>0.01</td>
<td>(−0.06 - 0.07)</td>
<td>0.87</td>
</tr>
<tr>
<td>Self-employed, conditional on working</td>
<td>−0.00</td>
<td>(−0.07 - 0.07)</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<sup>e</sup> Testosterone associations with outcomes were estimated separately, adjusting for confounders but not each other.

<sup>f</sup> Adjusted for age, age², time of day of testosterone sampling, % body fat, smoking status, GPS/BHPS subsample, beta-blockers, statins, NSAIDs, antidepressants, government office region and first ten principal components.

<sup>g</sup> Adjusted for age, age² and first ten principal components.

<sup>h</sup> Income measures are log-transformed and based on monthly income.
sample size limitations precluded their application here. Lastly, greater at the lower end of the testosterone range. Although methods in work suggests may also be socially patterned (Watts et al., 2017). As SNPs in restricted age range of 22 years suggest this did not affect conclusions. Anti-inflammatory medication use 1.05 (0.93–1.18) 0.47 3663 Antidepressant use 1.03 (0.92–1.16) 0.59 3663 Partnership (married or cohabiting vs other) 1.02 (0.94–1.10) 0.67 3663 Own children in household 1.00 (0.91–1.10) 0.98 3663

* PGS associations with confounders were estimated separately, adjusting for age, age², and the first ten principal components but not each other. Based on linear regression for continuous variables, logistic regressions for binary variables.

does not contribute greatly to social inequalities in health.

4.1. Strengths and limitations

To the author’s best knowledge, this is the first UK study to describe either age-differences or social differences in testosterone across the adult life-course. Using rich income data available in UKHLS, it was possible to examine patterning of testosterone by household income and by personal earnings, where different associations might be expected depending on the mechanisms involved, as well as self-employment and education. This is also the first study to apply Mendelian Randomization to testosterone and income, an association for which bias due to noise in single testosterone measures and substantial confounding by other factors is plausible. The difference between OLS and IV results indicate this approach is worthwhile. Meanwhile, this study had some limitations. Since measures of SHBG were not available, it was not possible to examine associations with free testosterone, which has potentially greater influence on social and health-related outcomes. However, total testosterone and free testosterone are very closely correlated in men younger than 70 (15, 16), and the similar results in the restricted age range of 22–70y suggest this did not affect conclusions. SNPs in SHBG could influence earnings via SHBG, which observational work suggests may also be socially patterned (Watts et al., 2017). As with any Mendelian Randomization, inflation of IV estimates by pleiotropy (an influence of the gene variants on SEP via pathways independent of testosterone) could not be ruled out. With only 3 genetic variants, traditional methods for assessing likely pleiotropy could not be used. However, the gene variants have not to date been associated in genome-wide association studies with any plausible confounders, and the polygenic score did not predict likely confounders in this sample. Risk-taking behaviour was self-assessed, and different associations may have been seen with a different or experimentally-derived measure. We do not know which men had previously had a vasectomy, which could influence circulating testosterone (Mo et al., 1995). Including an outlier for the PGS with very low testosterone strengthened associations with earnings and made that result conventionally statistically significant, suggesting impact of genetically-influenced testosterone on SEP may be greater at the lower end of the testosterone range. Although methods for nonlinear MR have been recently developed (Burgess et al., 2014), sample size limitations precluded their application here. Lastly, increasing evidence suggests associations of testosterone with behaviour may be modified by cortisol. Since cortisol was not measured in this survey, we could not examine interactive effects. As with any study using genetic variants explaining a modest proportion of variance in a trait, results will be subject to stochasticity and idiosyncrasies of the sample, and should be replicated in other study populations.

5. Conclusion

In a sample if British men aged 16–97, there was no evidence for a social gradient in circulating total testosterone by education or household income, nor of an association of testosterone with self-employment. Results of Mendelian Randomization models were consistent with a causal influence of testosterone on earnings and probability of being in work, although imprecision of IV estimates precludes drawing firm conclusions. The discrepancy suggests these aspects of SEP may be influenced more by lifetime average values of testosterone than short-term variation, or that confounding as well as measurement error may affect detectability of associations in observational studies. More work will be required to replicate these findings, and to explore how associations differ in women.

Acknowledgements

AH completed this work while at the Institute for Social and Economic Research, University of Essex, and supported by the Economic and Social Research Council grant ES/M008592/1. MK is supported by the University of Essex and ESRC grant RES-596-28-0001. We thank Laura Fumagalli for her valuable comments regarding interpretation of results.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2018.11.004.

References


