



Watts , E. L., Fensom , G. K., Smith Byrne , K., Perez-Cornago , A., Allen , N. E., Knuppel , A., Gunter , M. J., Holmes , M. V., Martin, R. M., Murphy, N., Tsilidis , K. K., Yeap , B. B., Key , T. J., & Travis , R. C. (2021). Circulating insulin-like growth factor-I, total and free testosterone concentrations and prostate cancer risk in 200,000 men in UK Biobank. *International Journal of Cancer*, 148(9), 2274-2288. <https://doi.org/10.1002/ijc.33416>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1002/ijc.33416](https://doi.org/10.1002/ijc.33416)

[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 Wiley at <https://doi.org/10.1002/ijc.33416>. 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/red/research-policy/pure/user-guides/ebr-terms/>

# Circulating insulin-like growth factor-I, total and free testosterone concentrations and prostate cancer risk in 200 000 men in UK Biobank

Eleanor L. Watts<sup>1</sup>  | Georgina K. Fensom<sup>1</sup>  | Karl Smith Byrne<sup>2</sup> |  
 Aurora Perez-Cornago<sup>1</sup>  | Naomi E. Allen<sup>3,4</sup> | Anika Knuppel<sup>1</sup>  |  
 Marc J. Gunter<sup>5</sup> | Michael V. Holmes<sup>3,6</sup>  | Richard M. Martin<sup>7,8,9</sup>  |  
 Neil Murphy<sup>5</sup>  | Konstantinos K. Tsilidis<sup>10,11</sup>  | Bu B. Yeap<sup>12,13</sup> |  
 Timothy J. Key<sup>1</sup> | Ruth C. Travis<sup>1</sup>

<sup>1</sup>Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>2</sup>Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France

<sup>3</sup>Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>4</sup>UK Biobank Ltd, Stockport, UK

<sup>5</sup>Section of Nutrition and Metabolism, International Agency for Research on Cancer, Lyon, France

<sup>6</sup>Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK

<sup>7</sup>MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

<sup>8</sup>Bristol Medical School, Department of Population Health Sciences, University of Bristol, Bristol, UK

<sup>9</sup>National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, UK

<sup>10</sup>Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

<sup>11</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

<sup>12</sup>Medical School, University of Western Australia, Perth, Australia

<sup>13</sup>Department of Endocrinology and Diabetes, Fiona Stanley Hospital, Perth, Australia

## Correspondence

Eleanor L. Watts, Cancer Epidemiology Unit, Nuffield Department of Population Health, Richard Doll Building, Old Road Campus, OX3 7LF, Oxford, UK.

## Funding information

British Heart Foundation, Grant/Award Number: FS/18/23/33512; Cancer Research UK, Grant/Award Numbers: C18281/A19169, C60192/A28516, C8221/A19170, C8221/A20986, C8221/A29017; Hellenic Republic, Ministry of Education; Medical Research Council, Grant/Award Numbers:

## Abstract

Insulin-like growth factor-I (IGF-I) and testosterone have been implicated in prostate cancer aetiology. Using data from a large prospective full-cohort with standardised assays and repeat blood measurements, and genetic data from an international consortium, we investigated the associations of circulating IGF-I, sex hormone-binding globulin (SHBG), and total and calculated free testosterone concentrations with prostate cancer incidence and mortality. For prospective analyses, risk was estimated using multivariable-adjusted Cox regression in 199 698 male UK Biobank participants. Hazard ratios (HRs) were corrected for regression dilution bias using repeat hormone measurements from a subsample. Two-sample Mendelian

**Abbreviations:** CI, confidence interval; GWAS, genome-wide association study; HR, hazard ratio; ICD-10, International Classification of Diseases Tenth Revision Code 10; IGF-I, insulin like growth factor-I; LD, linkage disequilibrium; MR, Mendelian randomisation; OR, odds ratio; PSA, prostate-specific antigen; SHBG, sex hormone-binding globulin; SNP, single nucleotide polymorphism.

Eleanor L. Watts, Georgina K. Fensom and Karl Smith Byrne contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of UICC.

MC\_UU\_12013/1, MC\_UU\_12013/2, MC\_UU\_12013/3; Nuffield Department of Population Health, Grant/Award Number: Early Career Research Fellowship; Wellcome Trust, Grant/Award Number: LEAP 205212/Z/16/Z

randomisation (MR) analysis of IGF-I and risk used genetic instruments identified from UK Biobank men and genetic outcome data from the PRACTICAL consortium (79 148 cases and 61 106 controls). We used *cis*- and all (*cis* and *trans*) SNP MR approaches. A total of 5402 men were diagnosed with and 295 died from prostate cancer (mean follow-up 6.9 years). Higher circulating IGF-I was associated with elevated prostate cancer diagnosis (HR per 5 nmol/L increment = 1.09, 95% CI 1.05-1.12) and mortality (HR per 5 nmol/L increment = 1.15, 1.02-1.29). MR analyses also supported the role of IGF-I in prostate cancer diagnosis (*cis*-MR odds ratio per 5 nmol/L increment = 1.34, 1.07-1.68). In observational analyses, higher free testosterone was associated with a higher risk of prostate cancer (HR per 50 pmol/L increment = 1.10, 1.05-1.15). Higher SHBG was associated with a lower risk (HR per 10 nmol/L increment = 0.95, 0.94-0.97), neither was associated with prostate cancer mortality. Total testosterone was not associated with prostate cancer. These findings implicate IGF-I and free testosterone in prostate cancer development and/or progression.

#### KEYWORDS

IGF-I, Mendelian randomisation, prospective analysis, prostate cancer, testosterone

## 1 | INTRODUCTION

Prostate cancer is the second most common cancer in men worldwide and a leading cause of cancer death.<sup>1</sup> Few potentially modifiable risk factors have been identified, but circulating hormone concentrations are thought to play a role in prostate cancer aetiology.<sup>2,3</sup>

Insulin-like growth factor-I (IGF-I) is involved in cell proliferation, differentiation and apoptosis, and prospective studies have shown a positive association of circulating IGF-I concentration with prostate cancer risk.<sup>4</sup> Less is known about its potential role in prostate cancer progression or mortality.<sup>5</sup>

Androgens are integral to the maintenance of normal prostate function.<sup>6</sup> In the circulation, testosterone is bound to sex hormone-binding globulin (SHBG) and albumin. Approximately 2% of total testosterone circulates unbound or “free” and is postulated to be biologically active.<sup>7</sup> Observational pooled analysis of individual participant data from prospective studies indicated that men with very low free testosterone may have a lower risk of prostate cancer,<sup>8</sup> and a recent Mendelian randomisation (MR) study supports a positive association between free testosterone concentration and prostate cancer diagnosis.<sup>9</sup> However, it is unclear whether circulating free testosterone concentration is associated with prostate cancer mortality.<sup>8,10</sup> Epidemiological studies have also reported an inverse association between prostate cancer risk and circulating SHBG,<sup>8</sup> although results from MR analyses are inconclusive.<sup>9</sup>

Previous risk estimates for prostate cancer in relation to hormone concentration have generally been based on data from nested case-control studies with a single blood draw at baseline. The UK Biobank study has standardised measurements of hormones from baseline blood samples collected in the whole cohort (500 000 participants) as well as repeat measurements of the hormones in a subset (20 000).

### What's new?

Testosterone, insulin-like growth factor-I (IGF-I), and sex hormone-binding globulin (SHBG) all have been associated with prostate-cancer risk. In this large, prospective study, the authors analyzed how these circulating hormones might impact mortality as well as risk. They found that men with higher IGF-I had a higher risk of both prostate-cancer diagnosis and mortality. Men with higher free testosterone had an increased risk of prostate cancer, while men with higher SHBG had a decreased risk. These results support the roles of IGF-I and testosterone in prostate cancer development.

In this article, we aimed to examine the associations of serum concentrations of IGF-I, SHBG, total and free testosterone with prostate cancer incidence and mortality, using observational data from UK Biobank. For IGF-I, we investigated potential causal associations of IGF-I with prostate cancer using MR analyses, with genetic data from UK Biobank and the PRACTICAL consortium (based on 79 000 prostate cancer cases and 61 000 controls). MR analyses of SHBG, total and free testosterone and prostate cancer risk using these datasets have recently been published.<sup>9</sup> MR uses germline genetic variants as proxies of putative risk factors and estimates their associations with disease. As germline genetic variants are fixed and randomly allocated at conception, this technique minimises the possibility of confounding and reverse causality, and is therefore considered a useful approach towards causal inference.<sup>11</sup> By using these two complementary approaches, we were able to robustly investigate associations and assess causation.

## 2 | PATIENTS AND METHODS

### 2.1 | UK Biobank observational analysis

#### 2.1.1 | Study design

UK Biobank is a prospective cohort with open access for public health research. Details of the study protocol and data collection are available online (<http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf>) and elsewhere.<sup>12,13</sup>

In brief, all participants were registered with the UK National Health Service (NHS) and lived within 40 km of one of the UK Biobank assessment centres. Approximately 9.2 million people were initially invited to participate. Overall, 503 317 men and women aged 40 to 69 years consented to join the cohort and attended one of 22 assessment centres throughout England, Wales and Scotland between 2006 and 2010, a participation rate of 5.5%.<sup>13</sup>

#### 2.1.2 | Baseline assessment

At the baseline assessment visit, participants provided information on a range of sociodemographic, physical, lifestyle and health-related factors via a self-completed touchscreen questionnaire and a computer-assisted personal interview.<sup>13</sup> Weight and height were measured at the assessment centre.<sup>13</sup>

#### 2.1.3 | Blood sampling and biomarker assays

At recruitment, blood sampling was successfully performed in 99.7% of the cohort. Blood was collected in a serum separator tube and shipped to the central processing laboratory in temperature-controlled boxes at 4°C,<sup>14</sup> then aliquoted and stored in a central working archive at -80°C.<sup>15</sup> Serum concentrations of circulating IGF-I, SHBG, testosterone and albumin were measured in all participants. IGF-I (DiaSorin Liaison XL), SHBG and testosterone (Beckman Coulter AU5800) were determined by chemiluminescent immunoassays. Albumin was measured by a colorimetric assay (Beckman Coulter AU5800). Average within-laboratory (total) coefficients of variation for low, medium and high internal quality control level samples for each biomarker ranged from 2.1% to 8.3%. Full details of the assay methods and quality assurance protocols are available online ([https://biobank.ndph.ox.ac.uk/showcase/docs/serum\\_biochemistry.pdf](https://biobank.ndph.ox.ac.uk/showcase/docs/serum_biochemistry.pdf)).

#### 2.1.4 | Free testosterone estimation

Free testosterone concentrations were estimated using a formula based on the law of mass action from measured total testosterone, SHBG and albumin concentrations.<sup>16,17</sup>

### 2.1.5 | Repeat assessment

Participants who lived within a 35 km radius were invited to attend a repeat assessment clinic at the UK Biobank Co-ordinating Centre in Stockport between August 2012 and June 2013. Repeat assessments were completed in 20 000 participants (9000 men) with a response rate of 21%.<sup>18</sup>

#### 2.1.6 | Participant follow-up

Cancer registration data were provided via record linkage to the NHS Central Register and obtained via NHS Digital, until the censoring date (31 March 2016 in England and Wales and 31 October 2015 in Scotland). Death data for England and Wales were provided by NHS Digital and for Scotland by the Information and Statistics Division (censoring dates 31 January 2018 in England and Wales, and 30 November 2016 in Scotland). In the analysis of incident prostate cancer, the endpoint was defined as the first diagnosis of prostate cancer, or prostate cancer mortality (primary or otherwise) (International Classification of Diseases Tenth revision code [ICD-10] C61<sup>19</sup>), whichever was recorded first. In the analysis of prostate cancer mortality, the endpoint was prostate cancer as the primary cause of death. Person-years were calculated from the date of recruitment to the date of the first cancer registration (excluding non-melanoma skin cancer [ICD-10 C44]), death or censoring date, whichever occurred first.

#### 2.1.7 | Exclusion criteria

Our analytical dataset included 199 698 men; we excluded 9871 men with prevalent cancer (except C44: non-melanoma skin cancer), 13 509 men who did not have blood data available or who had biomarker measurements that did not pass quality control procedures,<sup>20</sup> 1685 participants for whom it was not possible to determine genetic sex or who were identified as being genetically female, 2326 men who reported taking hormone medication at baseline and 758 men who had no body mass index (BMI) data.

#### 2.1.8 | Statistical analysis

Hazard ratios (HRs) and 95% confidence intervals (CIs) of prostate cancer diagnosis and mortality were estimated using Cox proportional hazards models, with age as the underlying time variable. Analyses were stratified by geographic area (10 UK regions) and age at recruitment (<45, 45-49, 50-54, 55-59, 60-64, ≥65 years), and adjusted for Townsend deprivation score (fifths, unknown [0.1%]), racial/ethnic group (white, mixed background, Asian, black, other and unknown [0.5%]), height (<170, ≥170- $<175$ , ≥175- $<180$ , ≥180 cm and unknown [0.1%]), lives with a wife or partner (no, yes), body mass index (BMI) (<25, ≥25- $<30$ , ≥30- $<35$ , ≥35 kg/m<sup>2</sup>), cigarette smoking (never, former, current light smoker (1- $<15$  cigarettes per day), current heavy smoker (≥15 cigarettes per day), current (number of cigarettes per day unknown) and smoking status unknown [0.6%]), alcohol

consumption (non-drinkers, <1-<10, ≥10-<20, ≥20 g ethanol/day, unknown [0.5%]), and self-reported diabetes (no, yes and unknown [0.5%]). Adjustment covariates were defined a priori based on previous analyses of UK Biobank data and categories were used to allow for nonlinear associations.<sup>21</sup>

Blood biomarker measurements were also available for up to 7776 men who attended a repeat assessment clinic a median of 4.4 years after first blood collection.<sup>18</sup> Measurement error and within person variability using single measures at baseline leads to underestimation of risk (ie, regression dilution bias),<sup>22</sup> to provide more precise and generalisable risk estimates, HRs for trend were estimated per absolute increase in usual hormone concentrations, with correction for regression dilution bias using the McMahon-Peto method.<sup>22,23</sup>

In the categorical analyses, biomarker measurements were categorised into fifths based on the distribution in the whole cohort and HRs were calculated relative to the lowest fifth of each blood parameter. The variance of the log risk in each group was calculated (from the variances and covariances of the log risk) and used to obtain group-specific 95% CIs, which enable comparisons across different exposure categories.<sup>24</sup>

The proportional hazards assumption was examined using time-varying covariates and Schoenfeld residuals and revealed no evidence of deviation.

### 2.1.9 | Subgroup analyses

Subgroup analyses for incident prostate cancer were examined by the following categories: age at diagnosis (≤65, >65 years), time from blood collection to diagnosis (≤4, >4 years), age at blood collection (<60, ≥60 years), BMI (<30, ≥30 kg/m<sup>2</sup>), smoking status (never or former, current), alcohol consumption (<10, ≥10 g ethanol/day), education status (no university degree, university degree), currently married/cohabiting (no, yes), Townsend index (<median, ≥median), ethnicity (white, non-white), height (≤175, >175 cm), diabetes (no, yes), family history of prostate cancer (no, yes), poor self-rated health (no, yes) and median observed hormone concentrations (<median, ≥median). Subgroup categories were chosen a priori on the basis of data distributions and previous analyses by this research group.<sup>4,8</sup> Heterogeneity in the associations for case-specific variables (ie, age at diagnosis and time from blood collection to diagnosis) was examined using stratified Cox models based on competing risks and comparing the risk coefficients and standard errors in the two subgroups, and testing with a  $\chi^2$  for heterogeneity. For non-case-specific factors, heterogeneity was assessed using a  $\chi^2$  interaction term. Heterogeneity in the associations with prostate cancer mortality was not tested due to the limited statistical power.

### 2.1.10 | Further analyses

We examined the association of IGF-I with incident prostate cancer after additional adjustment for concentrations of free testosterone and SHBG, and the associations of total and free testosterone and

SHBG after further adjustment for IGF-I (with adjustment for biomarkers categorised into fifths and unknown). As a further examination of trend, the categorical variable representing the fifths of the hormones was replaced with a continuous variable that was scored as 0, 0.25, 0.5, 0.75 and 1, such that a unit increase in this variable can be taken to represent an 80 percentile increase in concentrations to enable comparison across hormones and with previous pooled analyses.<sup>4,8</sup> Analyses with prostate cancer diagnosis were repeated after hormone concentrations were divided into tenths.

## 2.2 | Mendelian randomisation analyses

### 2.2.1 | Genetic associations with circulating IGF-I

Single nucleotide polymorphisms (SNPs) associated with circulating IGF-I ( $P < 5 \times 10^{-8}$  significance threshold) were identified from a publicly available genome wide association study (GWAS) from 167 174 male UK Biobank participants of European ancestry.<sup>25</sup> SNPs were pruned by a linkage disequilibrium (LD) threshold of  $r^2 < 0.001$ . UK Biobank genotyping details are reported elsewhere.<sup>26</sup>

GWAS results were partitioned into one primary *cis*-SNP instrument within the IGF-I gene region on chromosome 12 (rs5742653), and 121 additional *trans*-SNPs (SNPs associated with circulating IGF-I concentrations that are not located in this gene region). These *cis* and *trans*-SNPs together explained 6.3% of the variance in circulating concentrations of IGF-I.<sup>27</sup> SNP rs numbers, nearest gene and effect estimates are displayed in Supplementary Table S1.

### 2.2.2 | Genetic instruments for prostate cancer

We used summary statistics for SNP associations with prostate cancer risk that were generated from 79 148 prostate cancer cases and 61 106 controls of European ancestry from the PRACTICAL, CRUK, CAPS, BPC3 and PEGASUS consortia.<sup>28,29</sup> In brief, 44 825 prostate cancer cases and 27 904 controls were genotyped using OncoArray (<http://epi.grants.cancer.gov/oncoarray/>), and data were also available from several previous prostate cancer GWAS: UK stage 1 and stage 2; CaPS 1 and CaPS 2; BPC3; NCI PEGASUS; and iCOGS. Genotype information was imputed for all samples using the October 2014 release of the 1000 Genomes Project data as the reference panel. Odds ratios (ORs) and SEs were estimated using logistic regression and then meta-analysed using an inverse variance fixed-effect approach.

Where genetic instruments in the two datasets were not identical, we used HaploReg<sup>30</sup> to identify SNPs in linkage disequilibrium ( $r^2 > 0.8$ ) to use as proxies.

### 2.2.3 | Statistical analysis

We used a two-sample MR approach to estimate IGF-I associations with overall prostate cancer risk, using UK Biobank as our genetic instruments for IGF-I and PRACTICAL for genetic outcome analyses.

The MR estimation for IGF-I and prostate cancer was conducted by the Wald ratio using the *cis*-SNP (rs5742653). We also conducted analyses incorporating all 122 IGF-I associated SNPs using the inverse-variance weighted method, as well as weighted median and mode-based methods to reduce the influence of pleiotropy.<sup>31</sup> The MR pleiotropy residual sum and outlier (MR-PRESSO) test was used to investigate the role of outliers.<sup>32</sup> To further assess the potential presence of horizontal pleiotropy, we used Cochran's Q for heterogeneity<sup>31</sup> and the intercept from the MR-Egger method. Additionally, we used leave-one-out analyses to test the sensitivity of our results to single SNP effects. PhenoScanner was used to assess pleiotropy of the genetic instruments.<sup>33,34</sup>

All analyses were performed using Stata version 14.1 (Stata Corporation, College Station, TX) and R version 3.2.3. All tests of significance were two-sided, and *P* values <.05 were considered statistically

significant. MR analyses were performed using the *TwoSampleMR* R package.<sup>35</sup>

### 3 | RESULTS

#### 3.1 | UK Biobank observational analyses

After a mean follow-up of 6.9 years (SD = 1.3 years), 5402 (2.7%) men were diagnosed with prostate cancer and 295 died from the disease. Table 1 summarises the baseline characteristics of study participants. Mean age at recruitment was 56.5 years (SD = 8.2), and mean BMI was 27.8 kg/m<sup>2</sup>. 29% reported having had a prostate-specific antigen (PSA) test prior to baseline and 14% had a family history of

**TABLE 1** Baseline characteristics and blood data for all men and for men who developed prostate cancer in UK Biobank

	All men (N = 199 698)	Men who developed prostate cancer (N = 5402)
<b>Sociodemographic</b>		
Age at recruitment (years), mean (SD)	56.5 (8.19)	62.1 (5.26)
Most deprived quintile, % (N)	19.7 (39280)	15.8 (854)
Black ethnicity, % (N)	1.45 (2882)	2.18 (117)
Not in paid/self-employment, % (N)	38.3 (76448)	56.7 (3062)
Living with partner, % (N)	92.9 (152616)	95.5 (4281)
<b>Anthropometric, mean (SD)</b>		
Height (cm)	175.7 (6.84)	175.1 (6.68)
BMI (kg/m <sup>2</sup> )	27.8 (4.23)	27.5 (3.79)
<b>Lifestyle, % (N)</b>		
Current cigarette smokers	12.5 (24855)	9.35 (502)
Drinking alcohol ≥20 g per day	43.7 (86905)	42.7 (2298)
Low physical activity (0–10 METs per week)	28.1 (54408)	26.6 (1393)
<b>Health history, % (N)</b>		
Hypertension	52.2 (104099)	58.6 (3163)
Diabetes	6.82 (13545)	5.91 (318)
Poor self-rated health	4.81 (9547)	3.14 (169)
<b>Prostate-specific factors, % (N)</b>		
Ever had a PSA test	28.7 (54139)	46.5 (2396)
Family history of prostate cancer	13.6 (14925)	25.0 (703)
<b>Baseline blood measures, mean (SD)</b>		
IGF-I (nmol/L)	21.9 (5.52)	21.6 (5.23)
SHBG (nmol/L)	39.5 (16.6)	41.9 (16.0)
Total testosterone (nmol/L)	12.0 (3.65)	12.0 (3.53)
Free testosterone (pmol/L)	209 (59.5)	200 (54.5)
Albumin (g/L)	45.6 (2.61)	45.2 (2.53)

Abbreviations: BMI, body mass index; IGF-I, insulin-like growth factor-I; METs, metabolic equivalent of tasks; PSA, prostate-specific antigen; SHBG, sex hormone-binding globulin.

prostate cancer. Means and SDs for baseline biomarker measurements are displayed in Table 1. Regression dilution ratios ranged between 0.57 (free testosterone) and 0.80 (IGF-I) (Supplementary Table S2).

### 3.1.1 | Associations between hormone concentrations and prostate cancer risk

Serum IGF-I concentration was positively associated with prostate cancer incidence (HR per 5 nmol/L increment = 1.09, 95% CI 1.05-1.12;  $P_{\text{trend}} < .0001$ , Figure 1) and prostate cancer mortality (HR per 5 nmol/L increment = 1.15, 95% CI 1.02-1.29;  $P_{\text{trend}} = .03$ , Figure 2).

Serum SHBG concentration was inversely associated with prostate cancer incidence (HR per 10 nmol/L increment = 0.95, 95% CI 0.94-0.97;  $P_{\text{trend}} < .0001$ , Figure 1), but was not associated with prostate cancer mortality (Figure 2). Free testosterone was positively associated with prostate cancer incidence (HR per 50 pmol/L increment = 1.10, 95% CI 1.05-1.15;  $P_{\text{trend}} < .0001$ , Figure 1) but not with mortality. Total testosterone was not associated with prostate cancer incidence or mortality.

Risk estimates with and without adjustment for regression dilution bias are shown in Supplementary Table S2.

### 3.1.2 | Subgroup analyses

There was no evidence of heterogeneity in the associations of IGF-I with incident prostate cancer by any of the selected characteristics (Figure 3). There was some evidence that the inverse association between SHBG and prostate cancer incidence varied by IGF-I concentration ( $P_{\text{het}} = .02$ ); only men with lower concentrations of IGF-I (< the study median) had a reduced risk of prostate cancer (HR per 10 nmol/L increment in SHBG = 0.94, 95% CI 0.92-0.97, Figure 4).

There was no evidence of heterogeneity in the associations of total or free testosterone with prostate cancer incidence by any characteristics except for diabetes status. For free testosterone, there was evidence that the magnitude of the association with incident prostate cancer was greater in men who were diabetic at baseline (HR per 50 pmol/L increment = 1.19, 95% CI 1.10-1.29) than in those who were not (HR = 1.05, 95% CI 1.02-1.07;  $P_{\text{het}} = .004$ , Figure 5). Total testosterone was also associated with prostate cancer in men with Type II diabetes, but not in men without diabetes (Supplementary Figure S2).

### 3.1.3 | Further analyses

Associations with incident prostate cancer remained broadly similar when the associations were examined across tenths of the distributions (Supplementary Figure S1), and per 80 percentile increase (Supplementary Table S3). Minimally adjusted results are displayed in Supplementary Table S4. Mutual adjustment for

hormones did not materially affect the risk estimates (Supplementary Table S4).

## 3.2 | Mendelian randomisation

MR analysis using the *cis*-SNP found that IGF-I was significantly associated with a 34% increased prostate cancer risk per 5 nmol/L increment (95% CI 1.07-1.68;  $P = .01$ ) (Table 2; Supplementary Figure S3).

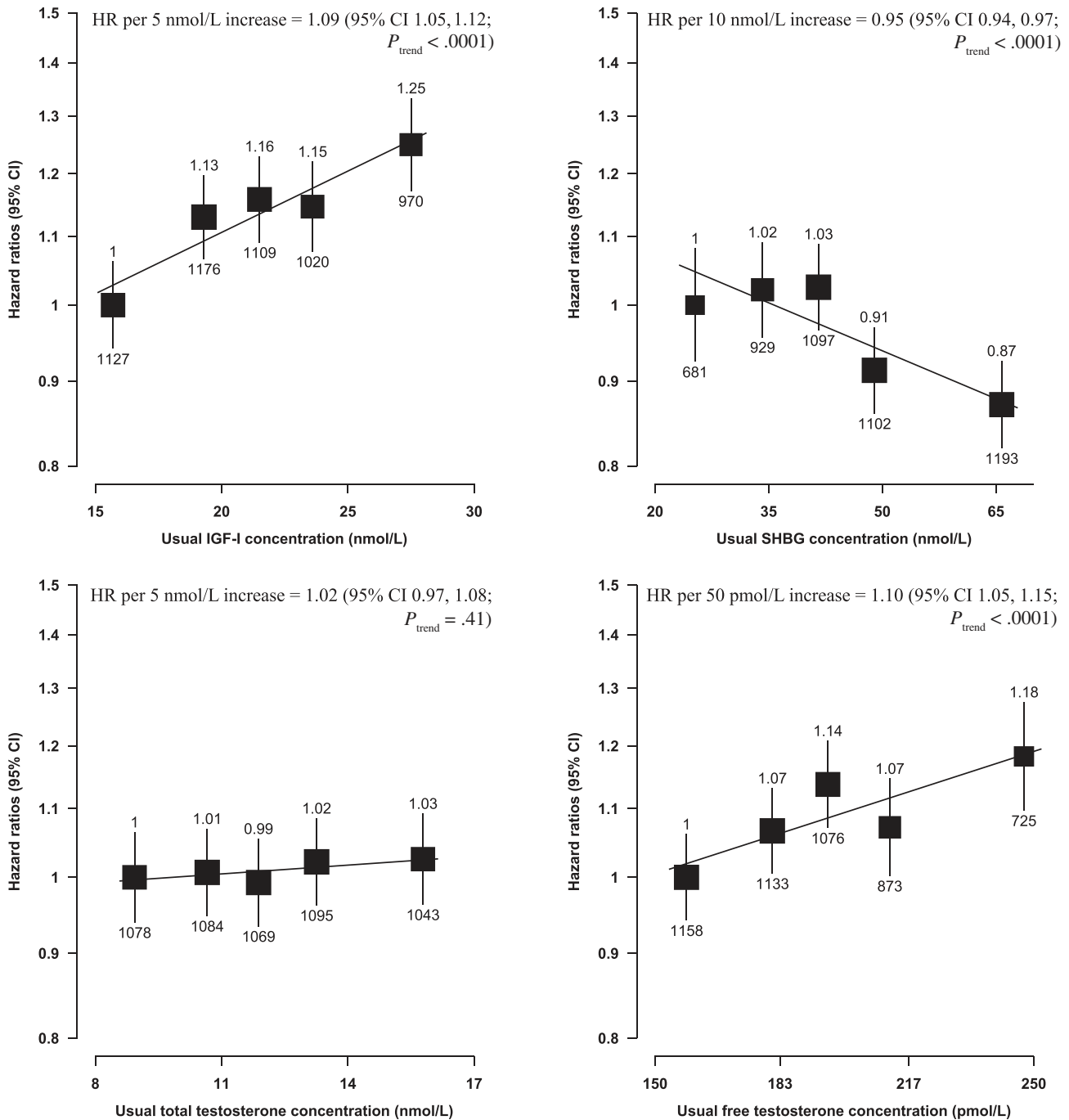
MR analysis including all the SNPs (both the *cis* and *trans*-SNPs) found suggestive evidence between IGF-I and prostate cancer diagnosis in the same direction as the *cis*-SNP results (inverse-variance weighted OR for a genetically predicted 5 nmol/L increment in IGF-I = 1.06, 95% CI 1.00-1.13;  $P = .06$ , Table 2). Although the Egger intercept did not indicate the presence of directional pleiotropy, significant heterogeneity (Cochran's  $Q$   $P < .0001$ ) may have influenced SE estimation for the inverse-variance weighted estimate (Table 2). Results were consistent following the removal of outliers identified using the MR-PRESSO outlier test (Table 2), and leave-one-out analyses, including removal of the *cis*-SNP (data not shown).

There is no direct functional evidence for the *cis*-SNP. However, this SNP is in modest linkage disequilibrium with the SNPs rs17727841, rs10860864, rs5742671, rs1996656 and rs11111250, which are expression quantitative trait loci (eQTLs) associated with *IGF1* and *WASHC3* expression.<sup>36</sup> PheWAS using published data showed the *cis*-SNP was associated with measures of lung function and adiposity (Supplementary Table S5). There was a large amount of pleiotropy in the *trans* genetic instruments, for example, for the top 100 *trans*-SNPs most strongly associated with IGF-I, there were >2500 published genome-wide significant associations (using the PheScanner resource).

## 4 | DISCUSSION

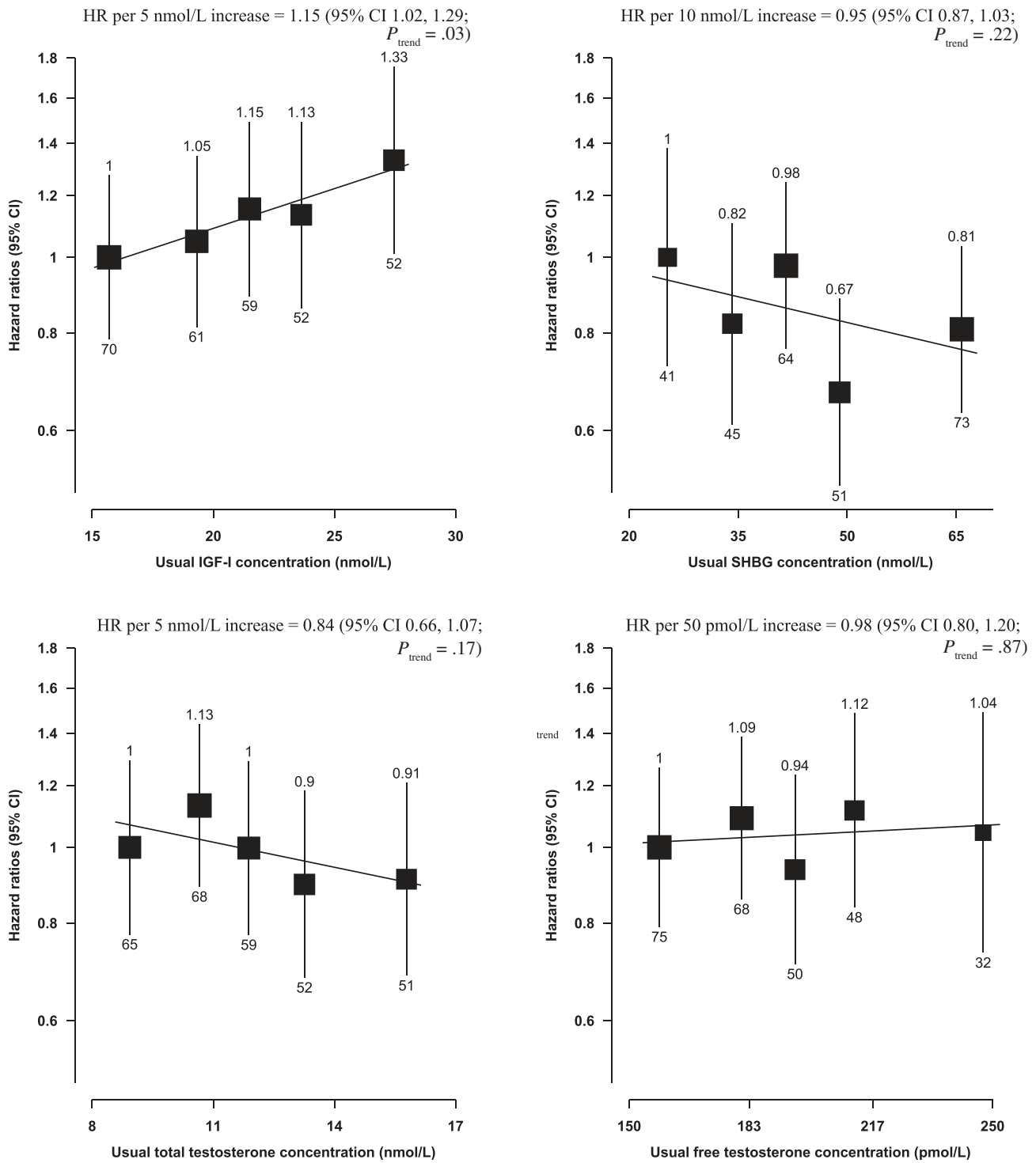
Our observational and MR analyses provide strong evidence that men with higher circulating IGF-I have an elevated risk of prostate cancer; furthermore, our observational analyses suggest a higher risk of prostate cancer mortality in these men, suggesting that IGF-I is associated with risk for more severe forms of prostate cancer and/or may increase the risk of prostate cancer progression. Higher serum-free testosterone was associated with a higher risk of prostate cancer diagnosis, which is supported by a recent MR analysis.<sup>9</sup> We also found that men with a higher SHBG had a lower risk. Total testosterone concentration was not associated with prostate cancer incidence or mortality.

The findings of a likely causal effect of IGF-I in prostate cancer development may be due to its role in activating signalling pathways, which regulate cell proliferation and apoptosis.<sup>2</sup> The positive relationship between IGF-I and incident prostate cancer observed is consistent with previous epidemiological evidence,<sup>4</sup> as well as associations observed with other cancers including breast and colorectal.<sup>37-39</sup> Further genetic epidemiology including fine mapping may help elucidate

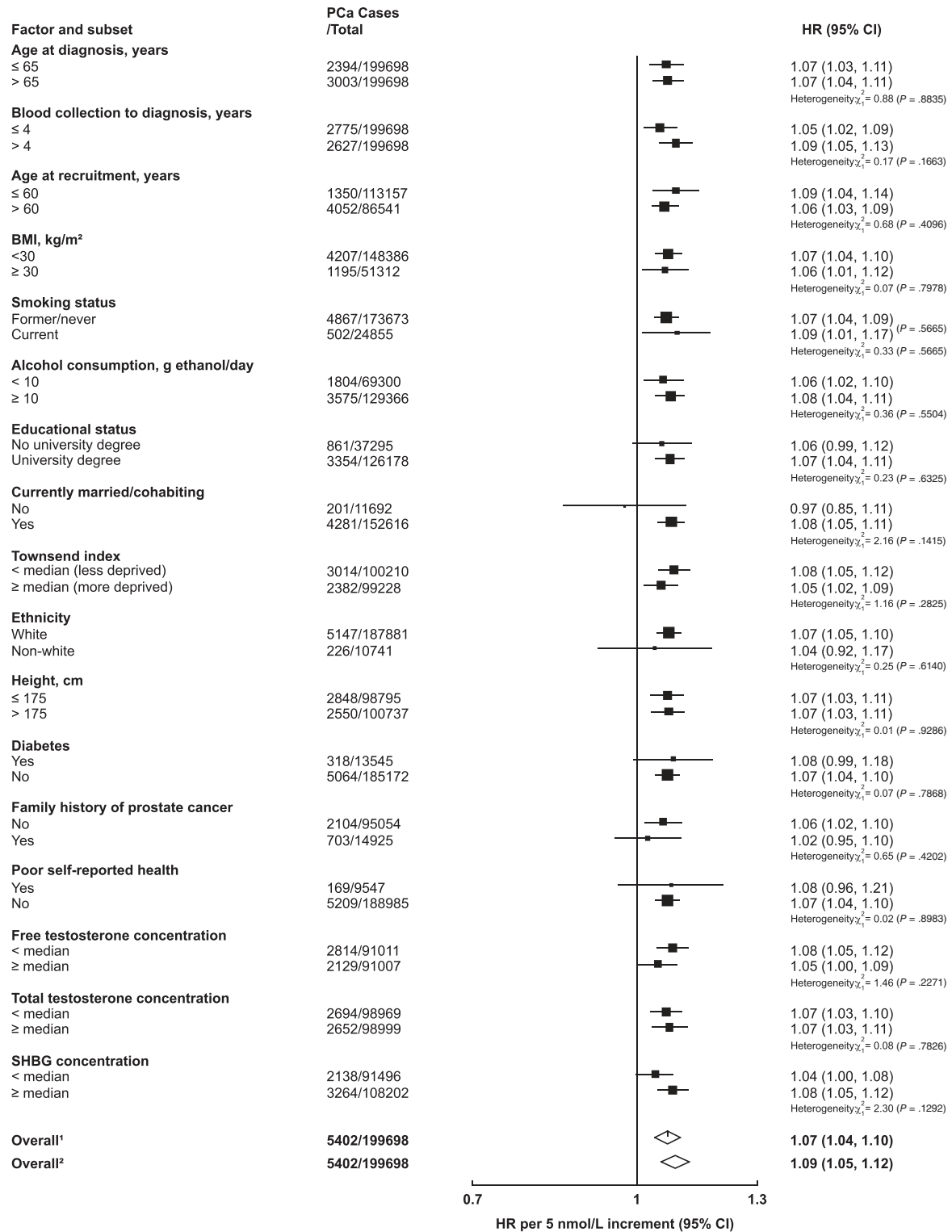


**FIGURE 1** Hazard ratios of incident prostate cancer by fifths of usual serum hormone concentrations in UK Biobank. HRs are stratified by region (10 UK cancer registry regions) and age at recruitment (<45, 45-49, 50-54, 55-59, 60-64 and ≥65 years) and adjusted for age (underlying time variable), Townsend deprivation score (fifths, unknown), racial/ethnic group (white, mixed background, Asian, black, other unknown), height (<170, ≥170-<175, ≥175-<180, ≥180 cm, unknown), lives with a wife or partner (no, yes), BMI (<25, ≥25-<30, ≥30-<35, ≥35 kg/m<sup>2</sup>), cigarette smoking (never, former, light smoker, heavy smoker, current unknown and smoking status unknown), alcohol consumption (non-drinkers, <1-<10, ≥10-<20, ≥20 g ethanol/day, unknown) and diabetes (no, yes and unknown). HRs for trend are adjusted for regression dilution bias. The boxes represent the HRs; the vertical lines represent the 95% CIs, with the size inversely proportional to the variance of the logarithm of the HR. The numbers above the vertical lines are point estimates for HRs, and the numbers below are the number of prostate cancer diagnoses. BMI, body mass index; CI, confidence intervals; HR, hazard ratio; IGF-I, insulin-like growth factor-I; SHBG, sex hormone-binding globulin

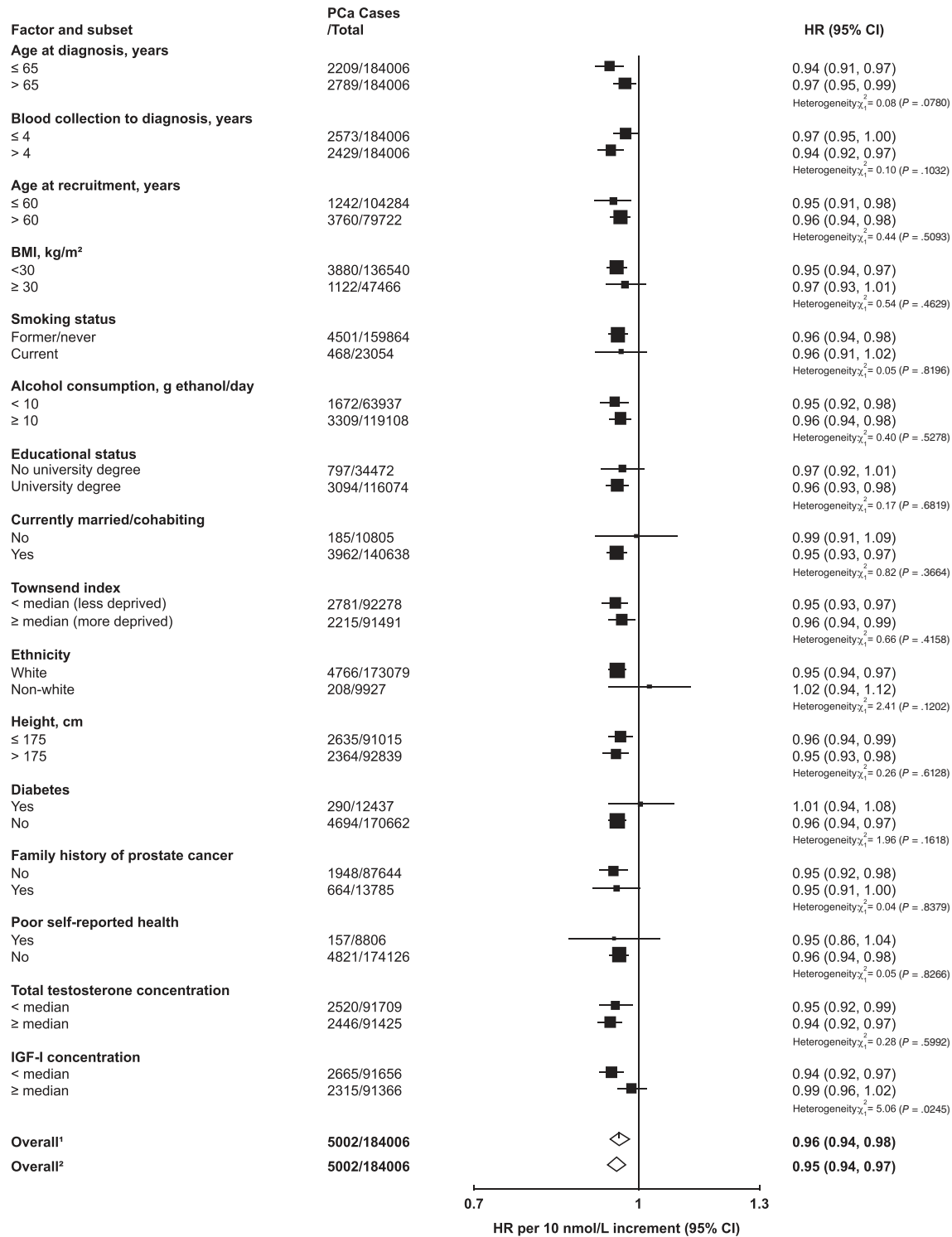




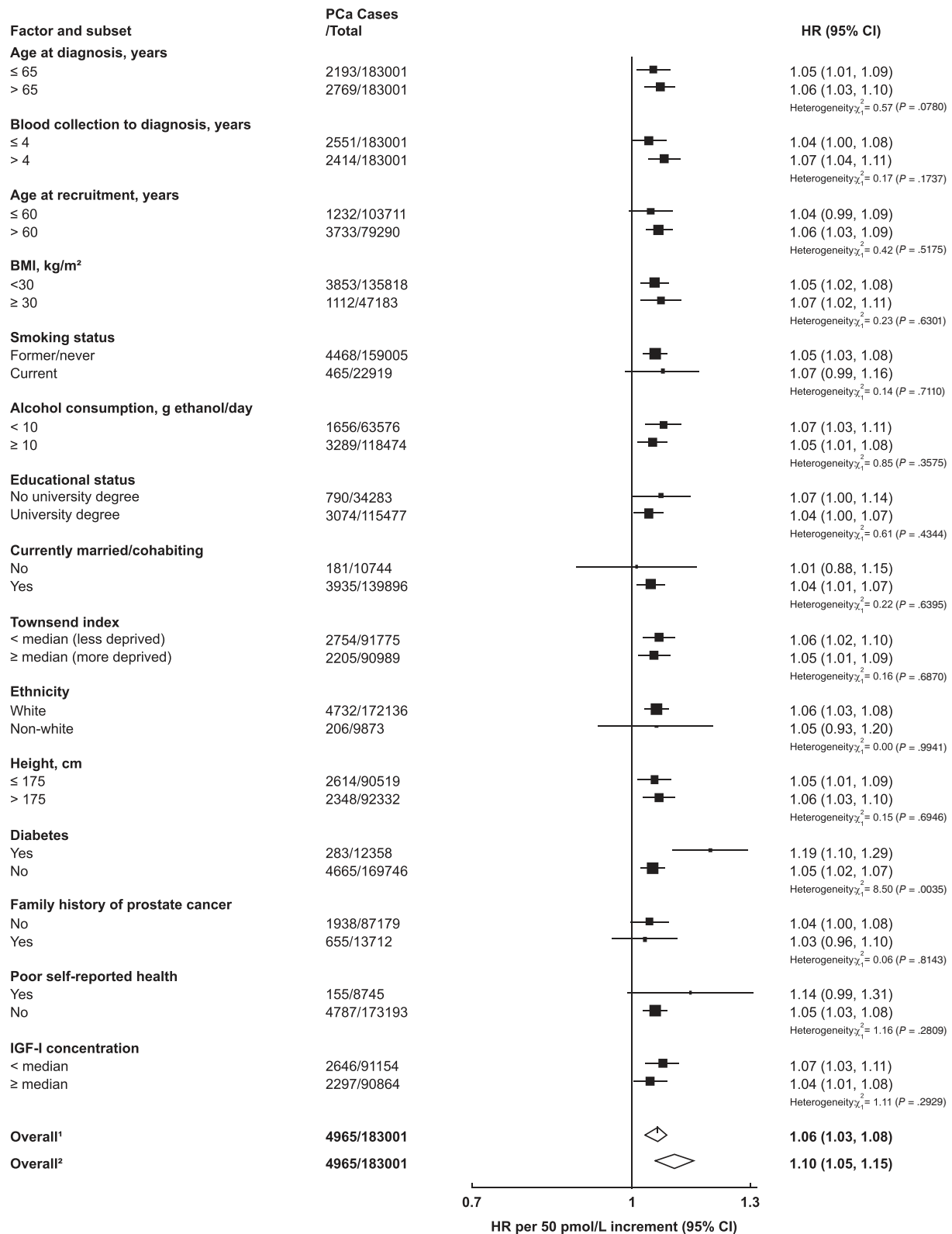
**FIGURE 2** Hazard ratios of prostate cancer mortality by fifths of usual serum hormone concentrations in the UK Biobank. HRs are stratified by region (10 UK cancer registry regions) and age at recruitment (<45, 45-49, 50-54, 55-59, 60-64, and ≥65 years) and adjusted for age (underlying time variable), Townsend deprivation score (fifths, unknown), racial/ethnic group (white, mixed background, Asian, black, other, unknown), height (<170, ≥170-<175, ≥175-<180, ≥180 cm, unknown), lives with a wife or partner (no, yes), BMI (<25, ≥25-<30, ≥30-<35, ≥35 kg/m<sup>2</sup>), cigarette smoking (never, former, light smoker, heavy smoker, current unknown and smoking status unknown), alcohol consumption (non-drinkers, <1-<10, ≥10-<20, ≥20 g ethanol/day, unknown) and diabetes (no, yes and unknown). HRs for trend are adjusted for regression dilution bias. The boxes represent the HRs; the vertical lines represent the 95% CIs, with the size inversely proportional to the variance of the logarithm of the HR. The numbers above the vertical lines are point estimates for HRs, and the numbers below are the number of prostate cancer deaths. BMI, body mass index; CI, confidence intervals; HR, hazard ratio; IGF-I, insulin-like growth factor-I; SHBG, sex hormone-binding globulin



**FIGURE 3** Hazard ratios of incident prostate cancer per 5 nmol/L increase in serum IGF-I concentration by subgroup in the UK Biobank. Cox models based on competing risks and compared the risk coefficients and SEs in the two subgroups and tested using a  $\chi^2$  test of heterogeneity. For non-case specific factors, heterogeneity was assessed using a  $\chi^2$  interaction term. HRs are stratified by region (10 UK cancer registry regions) and age at recruitment (<45, 45-49, 50-54, 55-59, 60-64, and ≥65 years) and adjusted for age (underlying time variable), Townsend deprivation score (fifths, unknown), racial/ethnic group (white, mixed background, Asian, black, other, unknown), height (<170, ≥170-175, ≥175-180, ≥180 cm, unknown), lives with a wife or partner (no, yes), BMI (<25, ≥25-30, ≥30-35, ≥35 kg/m<sup>2</sup>), cigarette smoking (never, former, light smoker, heavy smoker, current unknown and smoking status unknown), alcohol consumption (non-drinkers, <1-10, ≥10-20, ≥20 g ethanol/day, unknown) and diabetes (no, yes and unknown). The boxes represent the HRs; the horizontal lines represent the 95% CIs, with the size inversely proportional to the variance of the logarithm of the HR. (1) Not adjusted for regression dilution bias. (2) Adjusted for regression dilution bias. BMI, body mass index; CI, confidence intervals; HR, hazard ratio; IGF-I, insulin-like growth factor-I; PCa, prostate cancer



**FIGURE 4** Hazard ratio of incident prostate cancer per 10 nmol/L increase in serum SHBG concentration by subgroup in the UK Biobank. Cox models based on competing risks and compared the risk coefficients and SEs in the two subgroups and tested using a  $\chi^2$  test of heterogeneity. For non-case-specific factors, heterogeneity was assessed using a  $\chi^2$  interaction term. HRs are stratified by region (10 UK cancer registry regions) and age at recruitment (<45, 45-49, 50-54, 55-59, 60-64 and  $\geq 65$  years) and adjusted for age (underlying time variable), and adjusted for Townsend deprivation score (fifths, unknown), racial/ethnic group (white, mixed background, Asian, black, other, unknown), height (<170,  $\geq 170$ -<175,  $\geq 175$ -<180,  $\geq 180$  cm, unknown), lives with a wife or partner (no, yes), BMI (<25,  $\geq 25$ -<30,  $\geq 30$ -<35,  $\geq 35$  kg/m<sup>2</sup>), cigarette smoking (never, former, light smoker, heavy smoker, current unknown and smoking status unknown), alcohol consumption (non-drinkers, <1-<10,  $\geq 10$ -<20,  $\geq 20$  g ethanol/day, unknown) and diabetes (no, yes and unknown). The boxes represent the HRs; the horizontal lines represent the 95% CIs, with the size inversely proportional to the variance of the logarithm of the HR. (1) Not adjusted for regression dilution bias. (2) Adjusted for regression dilution bias. BMI, body mass index; CI, confidence intervals; HR, hazard ratio; PCa, prostate cancer; SHBG, sex hormone-binding globulin



**FIGURE 5** Hazard ratio of incident prostate cancer per 50 pmol/L increase in serum free testosterone concentration by subgroup in the UK Biobank. Cox models based on competing risks and compared the risk coefficients and SEs in the two subgroups and tested using a  $\chi^2$  test of heterogeneity. For non-case-specific factors, heterogeneity was assessed using a  $\chi^2$  interaction term. HRs are stratified by region (10 UK cancer registry regions) and age at recruitment (<45, 45-49, 50-54, 55-59, 60-64, and ≥65 years) and adjusted for age (underlying time variable), and adjusted for Townsend deprivation score (fifths, unknown), racial/ethnic group (white, mixed background, Asian, black, other, unknown), height (<170, ≥170-175, ≥175-180, ≥180 cm, unknown), lives with a wife or partner (no, yes), BMI (<25, ≥25-30, ≥30-35, ≥35 kg/m<sup>2</sup>), cigarette smoking (never, former, light smoker, heavy smoker, current unknown and smoking status unknown), alcohol consumption (non-drinkers, <1-10, ≥10-20, ≥20 g ethanol/day, unknown) and diabetes (no, yes and unknown). The boxes represent the HRs; the horizontal lines represent the 95% CIs, with the size inversely proportional to the variance of the logarithm of the HR. (1) Not adjusted for regression dilution bias. (2) Adjusted for regression dilution bias. BMI, body mass index; CI, confidence intervals; HR, hazard ratio; PCa, prostate cancer

**TABLE 2** Mendelian randomisation estimates between genetically predicted circulating IGF-I concentrations and prostate cancer risk

	Method	OR per genetically predicted 5 nmol/L increase in IGF-I (95% CI)	P value
<i>cis</i> -SNP (N = 1)	Wald ratio	1.34 (1.07–1.68)	.01
All SNPs ( <i>trans</i> - and <i>cis</i> -SNPs, N = 122)	Inverse weighted variance	1.06 (1.00–1.13)	.06
	Weighted median	1.03 (0.96–1.11)	.43
	Weighted mode	0.99 (0.87–1.11)	.82
	MR-Egger	1.04 (0.90–1.20)	.57
	MR-PRESSO	1.06 (1.00–1.12)	.04

Abbreviations: CI, confidence interval; IGF-I, insulin-like growth factor-I; OR, odds ratio; SNP, single nucleotide polymorphism.

exactly by which mechanism variation at the *IGF1* locus associates with risk of prostate and some other cancers.

Our finding of a positive association between calculated free testosterone concentration and prostate cancer diagnosis is consistent with the importance of androgens for prostate cancer development,<sup>3</sup> previous evidence from pooled nested case-control studies,<sup>8</sup> randomised controlled trials that aim to reduce intraprostatic androgen signalling<sup>40,41</sup> and MR analyses.<sup>9</sup> However, the shape of the association between circulating free testosterone and prostate cancer risk is not completely clear; here we see an approximately linear association, whereas our previous pooled analysis suggested that very low free testosterone concentrations were associated with a lower prostate cancer risk, but that risk did not change with further increments in free testosterone concentration.<sup>8</sup>

There is some evidence to suggest that men with low free testosterone may have an increased risk of high-grade prostate tumours, but the data are inconclusive.<sup>8,40–42</sup> In this analysis, prostate cancer mortality was used as a proxy of tumour aggressiveness; we did not observe an association between circulating total or free testosterone concentrations and prostate cancer mortality, although statistical power to examine this association was limited (<300 prostate cancer deaths).

Men with higher SHBG concentrations had a lower risk of prostate cancer, which is consistent with previous prospective studies.<sup>8</sup> MR analyses have also shown some evidence of an inverse relationship, but results were not statistically significant.<sup>9</sup> In the current study, we cannot determine whether the mechanism underlying this association relates to SHBG itself or to the action of SHBG as a carrier protein, which modulates androgen access to tissues.

This analysis has several strengths. It is the largest prospective full-cohort analysis to examine hormones in relation to prostate cancer incidence and mortality. UK Biobank is a well-characterised study population, and therefore we were able to adjust our risk results for a wide range of possible confounders and to investigate associations in a number of subgroups. Our results were consistent across these subgroups; in only three subgroup analyses, there was weak evidence of heterogeneity in the associations with prostate cancer risk. We observed heterogeneity in the associations of total and free testosterone by diabetes status. This heterogeneity may be related to hormone differences in diabetic men, who have a lower risk of prostate cancer

diagnosis.<sup>21</sup> However, men with diabetes also have lower PSA concentrations, which reduces the probability of prostate cancer detection<sup>43,44</sup> and may lead to differential bias of our risk estimates; we are also unable to rule-out chance due to multiple testing and the small numbers of men with diabetes and subsequent prostate cancer (n = 283–311). Hormones were measured using a standardised method; therefore, we were able to estimate risk associations of the absolute scale and use repeat measurements to improve the precision of risk estimates.<sup>22</sup> Furthermore, by incorporating both observational and MR methods, we were able to use different lines of evidence with orthogonal biases to investigate the potential causality of the associations of IGF-I with prostate cancer risk.<sup>45</sup> Our MR analysis of IGF-I using a *cis*-SNP is an example of the strongest case for an MR analysis, due to the strong plausibility of a biological link and a reduced likelihood of horizontal pleiotropy<sup>46,47</sup>; therefore the association for this *cis*-SNP indicates that IGF-I may be driving the reported associations with prostate cancer risk.

A limitation of the analysis is that prostate tumour stage and grade information are not currently available in the UK Biobank, and only incident prostate cancer data are available in the MR analyses. The UK Biobank participants are predominantly white and healthier than the sampling population; therefore, selection bias may influence the results<sup>48</sup> and risk estimates may not be generalisable,<sup>13</sup> although this is unlikely to affect the direction of the associations.<sup>49</sup> Relatively weak evidence from MR analyses incorporating all GWAS significant SNPs for IGF-I may reflect widespread pleiotropy for *trans*-SNPs including the IGF1 and IGF-II signalling, which served as primary motivation to emphasise the association of the *cis*-variant. A limitation of the use of this *cis* instrument is the lack of direct functional evidence; however, the identification of an individual SNP linked to an eQTL is difficult because of the high collinearity of SNPs. There was some evidence that SNPs in modest linkage disequilibrium with the *cis*-variant alter the expression levels of *IGF1* and *WASHC3*. Due to the strong association of rs5742653 with circulating IGF-I concentrations, its location within the *IGF1* gene region, and the lack of any other biological evidence linking *WASHC3* to prostate cancer risk, we consider that the use of this *cis*-SNP is valid. The effect estimates from the MR analyses were calculated on the same scale as for the observational analyses, and this scaling-up results in some imprecision with wide confidence intervals in the association with the *cis*-SNP;

the concordance of the directions of the associations is particularly important.

Testosterone is related to other androgens, which have not been measured in UK Biobank; therefore, associations may be at least partially explained by other androgens, although there is little observational evidence to support this.<sup>50</sup> Free testosterone concentrations were estimated using a commonly employed formula derived from mass action equations, and the results may not correspond precisely to free testosterone measured using equilibrium dialysis.<sup>51</sup> SHBG circulates as a homodimer, which presents two binding sites for testosterone adding further complexity to calculation of free testosterone.<sup>7</sup> Furthermore, experimental evidence postulates entry of SHBG-bound testosterone into cells via megalin, an endocytic receptor expressed in reproductive tissues.<sup>52</sup> Therefore, the predictive value of calculated free testosterone as an indicator of the bioavailability of testosterone to the cells within the prostate remains under debate.<sup>53</sup> Measurement of serum-free testosterone using equilibrium dialysis may improve the accuracy of risk estimates, but this methodology is labour-intensive and not readily available.

PSA concentrations are partly regulated by the androgen receptor<sup>54</sup>; lower free testosterone concentrations may therefore reduce circulating PSA concentrations, reducing the likelihood of prostate cancer detection (possibly as well as reducing the likelihood of cancer development). Comorbidities, socioeconomic status and poor health may affect PSA test attendance, but PSA testing attendance after baseline was not known.

In conclusion, our results implicate IGF-I and free testosterone in prostate cancer development and/or progression. This analysis of 200 000 men enabled us to quantify the associations of circulating hormone concentrations with prostate cancer risk. The complementary MR for IGF-I supports a causal association. Future research will examine hormone associations by tumour stage and grade.

## ACKNOWLEDGEMENTS

This research has been conducted using the UK Biobank Resource under application number 3282 and data from the PRACTICAL consortium, CRUK, BPC3, CAPS and PEGASUS. This work was supported by Cancer Research UK (grant numbers C8221/A19170, C8221/A20986 and C8221/A29017). ELW was supported by the Nuffield Department of Population Health Early Career Research Fellowship. AK is supported by the Wellcome Trust (LEAP 205212/Z/16/Z). APC is supported by a Cancer Research UK Population Research Fellowship (C60192/A28516). MVH works in a unit that receives funding from the UK Medical Research Council and is supported by a British Heart Foundation Intermediate Clinical Research Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre. RMM was supported by a Cancer Research UK (C18281/A19169) programme grant (the Integrative Cancer Epidemiology Programme) and is part of the Medical Research Council Integrative Epidemiology Unit at the University of Bristol supported by the Medical Research Council (MC\_UU\_12013/1, MC\_UU\_12013/2, and MC\_UU\_12013/3) and the University of Bristol. RMM is also supported by the National Institute for Health

Research (NIHR) Bristol Biomedical Research Centre, which is funded by the National Institute for Health Research (NIHR) and is a partnership between University Hospitals Bristol NHS Foundation Trust and the University of Bristol. KKT was funded by the Small Research Teams funding programme from the Hellenic Republic, Ministry of Education.

The Prostate cancer genome-wide association analyses are supported by the Canadian Institutes of Health Research, European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175), Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A16563, C5047/A3354, C5047/A10692, C16913/A6135 and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative).

We would also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now PCUK), The Orchid Cancer Appeal, Rosetrees Trust, The National Cancer Research Network UK and The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust.

The Prostate Cancer Program of Cancer Council Victoria also acknowledge grant support from The National Health and Medical Research Council, Australia (126402, 209057, 251533, 396414, 450104, 504700, 504702, 504715, 623204, 940394, 614296), VicHealth, Cancer Council Victoria, The Prostate Cancer Foundation of Australia, The Whitten Foundation, PricewaterhouseCoopers and Tattersall's. EAO, DMK and EMK acknowledge the Intramural Program of the National Human Genome Research Institute for their support.

Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) [U19 CA 148537 for ELucidating Loci Involved in Prostate cancer Susceptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008] and by Cancer Research UK grant A8197/A16565. Additional analytic support was provided by NIH NCI U01 CA188392 (PI: Schumacher).

Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112—the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund.

The BPC3 was supported by the U.S. National Institutes of Health, National Cancer Institute (cooperative agreements U01-CA98233 to D.J.H., U01-CA98710 to S.M.G., U01-CA98216 to

E.R., and U01-CA98758 to B.E.H., and Intramural Research Program of NIH/National Cancer Institute, Division of Cancer Epidemiology and Genetics).

CAPS GWAS study was supported by the Swedish Cancer Foundation (Grant No 09-0677, 11-484, 12-823), the Cancer Risk Prediction Center (CRiSP; [www.crispcenter.org](http://www.crispcenter.org)), a Linneus Centre (Contract ID 70867902) financed by the Swedish Research Council, Swedish Research Council (Grant No K2010-70X-20430-04-3, 2014-2269).

PEGASUS was supported by the Intramural Research Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health.

## DISCLAIMERS

Department of Health and Social Care disclaimer: The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

## CONFLICT OF INTEREST

Dr Holmes has collaborated with Boehringer Ingelheim in research, and in adherence to the University of Oxford's Clinical Trial Service Unit & Epidemiological Studies Unit (CSTU) staff policy, did not accept personal honoraria or other payments from pharmaceutical companies. All other authors have no competing interests to declare.

## DATA AVAILABILITY STATEMENT

UK Biobank is an open access resource, and the study website <https://www.ukbiobank.ac.uk/> has information on available data and access procedures. GWAS results are publicly available from <http://www.nealelab.is/uk-biobank>. PRACTICAL data are available upon application via the consortium website <http://practical.icr.ac.uk/blog/>. Additional datasets used in the analyses will be made available upon reasonable request.

## ETHICS STATEMENT

The UK Biobank study was approved by the North West Multi-Centre Research Ethics Committee (reference number 06/MRE08/65) and at recruitment all participants gave written informed consent to participate and for their health to be followed-up through linkage to electronic medical records.

## ORCID

Eleanor L. Watts  <https://orcid.org/0000-0001-9229-2589>  
 Georgina K. Fensom  <https://orcid.org/0000-0003-0979-4436>  
 Aurora Perez-Cornago  <https://orcid.org/0000-0002-5652-356X>  
 Anika Knuppel  <https://orcid.org/0000-0003-1049-4836>  
 Michael V. Holmes  <https://orcid.org/0000-0001-6617-0879>  
 Richard M. Martin  <https://orcid.org/0000-0002-7992-7719>

Neil Murphy  <https://orcid.org/0000-0003-3347-8249>

Konstantinos K. Tsilidis  <https://orcid.org/0000-0002-8452-8472>

## REFERENCES

1. Ferlay J, Ervik M, Lam F, et al. *Global Cancer Observatory: Cancer Today*. International Agency for Research on Cancer: Lyon, France; 2018.
2. Pollak M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nat Rev Cancer*. 2012;12:159-169.
3. Morgentaler A, Traish AM. Shifting the paradigm of testosterone and prostate cancer: the saturation model and the limits of androgen-dependent growth. *Eur Urol*. 2009;55:310-320.
4. Travis RC, Appleby PN, Martin RM, et al. A meta-analysis of individual participant data reveals an association between circulating levels of IGF-I and prostate cancer risk. *Cancer Res*. 2016;76:2288-2300.
5. Major JM, Laughlin GA, Kritiz-Silverstein D, Wingard DL, Barrett-Connor E. Insulin-like growth factor-I and cancer mortality in older men. *J Clin Endocrinol Metab*. 2010;95:1054-1059.
6. Dai C, Heemers H, Sharifi N. Androgen signaling in prostate cancer. *Cold Spring Harb Perspect Med*. 2017;7:a030452.
7. Goldman AL, Bhasin S, Wu FCW, Krishna M, Matsumoto AM, Jasuja A. A reappraisal of testosterone's binding in circulation: physiological and clinical implications. *Endocr Rev*. 2017;38:302-324.
8. Watts EL, Appleby PN, Perez-Cornago A, et al. Low free testosterone and prostate cancer risk: a collaborative analysis of 20 prospective studies. *Eur Urol*. 2018;74:585-594.
9. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med*. 2020;26:252-258.
10. Claps M, Petrelli F, Caffo O, et al. Testosterone levels and prostate cancer prognosis: systematic review and meta-analysis. *Clin Genitourin Cancer*. 2018;16:165-75.e2.
11. Yarmolinsky J, Wade KH, Richmond RC, et al. Causal inference in cancer epidemiology: what is the role of Mendelian randomization? *Cancer Epidemiol Biomarkers Prev*. 2018;27:995-1010.
12. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12:e1001779.
13. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*. 2017;186:1026-1034.
14. Elliott P, Peakman TC. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int J Epidemiol*. 2008;37:234-244.
15. Woods P, Owen JM. Designing and implementing a large-scale automated  $-80^{\circ}\text{C}$  archive. *Int J Epidemiol*. 2008;37:i56-i61.
16. Bartsch W. Interrelationships between sex hormone-binding globulin and testosterone, 5 alpha-dihydrotestosterone and oestradiol-17 beta in blood of normal men. *Maturitas*. 1980;2:109-118.
17. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab*. 1999;84:3666-3672.
18. UK Biobank. Repeat Assessment Data, 2013 [https://biobank.ctsu.ox.ac.uk/~bbdatan/Repeat\\_assessment\\_doc\\_v1.0.pdf](https://biobank.ctsu.ox.ac.uk/~bbdatan/Repeat_assessment_doc_v1.0.pdf). Accessed December 4, 2020.
19. WHO. International Statistical Classification of Diseases and Related Health Problems. 10th revision. 2010.
20. UK Biobank. Biomarker Assay Quality Procedures: Approaches Used to Minimise Systematic and Random Errors (and the Wider Epidemiological Implications), 2019 [http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/biomarker\\_issues.pdf](http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/biomarker_issues.pdf). Accessed December 4, 2020.
21. Perez-Cornago A, Key TJ, Allen NE, et al. Prospective investigation of risk factors for prostate cancer in the UK Biobank cohort study. *Br J Cancer*. 2017;117:1562-1571.

22. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol.* 1999;150:341-353.
23. MacMahon S, Peto R, Cutler J, et al. Blood pressure, stroke, and coronary heart disease. Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet.* 1990;335:765-774.
24. Plummer M. Improved estimates of floating absolute risk. *Stat Med.* 2004;23:93-104. Accessed December 4, 2020.
25. NealeLab UK Biobank GWAS Results, 2019 <http://www.nealelab.is/uk-biobank>. Accessed December 4, 2020.
26. Bycroft Cea. Genome-Wide Genetic Data on ~500,000 UK Biobank Participants 2017.
27. Goddard ME, Wray NR, Verbyla K, Visscher PM. Estimating effects and making predictions from genome-wide marker data. *Stat Sci.* 2009;24:517-529.
28. Schumacher FR, Al Olama AA, Berndt SI, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet.* 2018;50:928-936.
29. Wu L, Wang J, Cai Q, et al. Identification of novel susceptibility loci and genes for prostate cancer risk: a transcriptome-wide association study in over 140,000 European descendants. *Cancer Res.* 2019;79:3192-3204.
30. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* 2016;44:D877-D881.
31. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: a review. *Res Synth Methods.* 2019;10:486-496.
32. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet.* 2018;50:693-698.
33. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics.* 2019;35:4851-4853.
34. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics.* 2016;32:3207-3209.
35. Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human phenome. *Elife.* 2018;7:e34408.
36. Zheng Z, Huang D, Wang J, et al. QTLbase: an integrative resource for quantitative trait loci across multiple human molecular phenotypes. *Nucleic Acids Res.* 2020;48:D983-D991.
37. Murphy N, Knuppel A, Papadimitriou N, et al. Insulin-like growth factor-1, insulin-like growth factor-binding protein-3, and breast cancer risk: observational and Mendelian randomization analyses with ~430 000 women. *Ann Oncol.* 2020;31:641-649.
38. Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. *Lancet Oncol.* 2010;11:530-542.
39. Murphy N, Carreras-Torres R, Song M, et al. Circulating levels of insulin-like growth factor 1 and insulin-like growth factor binding protein 3 associate with risk of colorectal cancer based on serologic and Mendelian randomization analyses. *Gastroenterology.* 2020;158:1300-1312.
40. Andriole GL, Bostwick DG, Brawley OW, et al. Effect of dutasteride on the risk of prostate cancer. *N Engl J Med.* 2010;362:1192-1202.
41. Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. *N Engl J Med.* 2003;349:215-224.
42. Platz EA, Leitzmann MF, Rifai N, et al. Sex steroid hormones and the androgen receptor gene CAG repeat and subsequent risk of prostate cancer in the prostate-specific antigen era. *Cancer Epidemiol Biomarkers Prev.* 2005;14:1262-1269.
43. Wu C, Moreira DM, Gerber L, Rittmaster RS, Andriole GL, Freedland SJ. Diabetes and prostate cancer risk in the REDUCE trial. *Prostate Cancer Prostatic Dis.* 2011;14:326-331.
44. Miller EA, Pinsky PF, Pierre-Victor D. The relationship between diabetes, prostate-specific antigen screening tests, and prostate cancer. *Cancer Causes Control.* 2018;29:907-914.
45. Lawlor DA, Tilling K, Davey SG. Triangulation in aetiological epidemiology. *Int J Epidemiol.* 2017;45:1866-1886.
46. Swerdlow DI, Kuchenbaecker KB, Shah S, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int J Epidemiol.* 2016;45:1600-1616.
47. Stacey D, Fauman EB, Ziemek D, et al. ProGeM: a framework for the prioritization of candidate causal genes at molecular quantitative trait loci. *Nucleic Acids Res.* 2018;47(1):e3.
48. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey SG. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol.* 2018;47:226-235.
49. Batty GD, Gale CR, Kivimäki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ.* 2020;368:m131.
50. Roddam AW, Allen NE, Appleby P, Key TJ. Endogenous sex hormones and prostate cancer: a collaborative analysis of 18 prospective studies. *J Natl Cancer Inst.* 2008;100:170-183.
51. Ly LP, Sartorius G, Hull L, et al. Accuracy of calculated free testosterone formulae in men. *Clin Endocrinol (Oxf).* 2010;73:382-388.
52. Hammes A, Andreassen TK, Spoelgen R, et al. Role of endocytosis in cellular uptake of sex steroids. *Cell.* 2005;122:751-762.
53. Handelsman DJ. Free testosterone: pumping up the tires or ending the free ride? *Endocr Rev.* 2017;38:297-301.
54. Kim J, Coetzee GA. Prostate specific antigen gene regulation by androgen receptor. *J Cell Biochem.* 2004;93:233-241.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Watts EL, Fensom GK, Smith Byrne K, et al. Circulating insulin-like growth factor-I, total and free testosterone concentrations and prostate cancer risk in 200 000 men in UK Biobank. *Int. J. Cancer.* 2021;148:2274-2288. <https://doi.org/10.1002/ijc.33416>