



**HAL**  
open science

## **Circulating insulin-like growth factors and risks of overall, aggressive and early-onset prostate cancer: a collaborative analysis of 20 prospective studies and Mendelian randomization analysis**

Eleanor Watts, Aurora Perez-Cornago, Georgina Fensom, Karl Smith-Byrne, Urwah Noor, Colm Andrews, Marc Gunter, Michael Holmes, Richard Martin, Konstantinos Tsilidis, et al.

### ► **To cite this version:**

Eleanor Watts, Aurora Perez-Cornago, Georgina Fensom, Karl Smith-Byrne, Urwah Noor, et al.. Circulating insulin-like growth factors and risks of overall, aggressive and early-onset prostate cancer: a collaborative analysis of 20 prospective studies and Mendelian randomization analysis. *International Journal of Epidemiology*, 2023, 52 (1), pp.71-86. 10.1093/ije/dyac124 . hal-04454915

**HAL Id: hal-04454915**

**<https://hal.science/hal-04454915>**

Submitted on 23 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Watts, E. L., Perez-Cornago, A., Fensom , G. K., Noor, U., Andrews, C. D., Smith-Byrne , K., Martin, R. M., Holly, J. M. P., the PRACTICAL Consortium, & et al. (2022). Circulating insulin-like growth factors and risks of overall, aggressive and early-onset prostate cancer: a collaborative analysis of 20 prospective studies and Mendelian randomization analysis. *International Journal of Epidemiology*, Article dyac124. <https://doi.org/10.1093/ije/dyac124>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1093/ije/dyac124](https://doi.org/10.1093/ije/dyac124)

[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 Oxford University Press at <https://doi.org/10.1093/ije/dyac124>. 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/>

Original article

# Circulating insulin-like growth factors and risks of overall, aggressive and early-onset prostate cancer: a collaborative analysis of 20 prospective studies and Mendelian randomization analysis

Eleanor L Watts <sup>1\*</sup>, Aurora Perez-Cornago <sup>1</sup>, Georgina K Fensom,<sup>1</sup> Karl Smith-Byrne,<sup>2</sup> Urwah Noor,<sup>1</sup> Colm D Andrews,<sup>1</sup> Marc J Gunter,<sup>3</sup> Michael V Holmes <sup>4,5</sup>, Richard M Martin,<sup>6,7,8</sup> Konstantinos K Tsilidis,<sup>9,10</sup> Demetrius Albanes,<sup>11</sup> Aurelio Barricarte,<sup>12,13,14</sup> H Bas Bueno-de-Mesquita,<sup>15</sup> Barbara A Cohn,<sup>16</sup> Melanie Deschasaux-Tanguy <sup>17</sup>, Niki L Dimou,<sup>3</sup> Luigi Ferrucci,<sup>18</sup> Leon Flicker <sup>19,20</sup>, Neal D Freedman,<sup>11</sup> Graham G Giles <sup>21,22,23</sup>, Edward L Giovannucci <sup>24,25,26</sup>, Christopher A Haiman,<sup>27</sup> Graham J Hankey,<sup>19</sup> Jeffrey MP Holly,<sup>28</sup> Jiaqi Huang <sup>11,29</sup>, Wen-Yi Huang <sup>11</sup>, Lauren M Hurwitz,<sup>11</sup> Rudolf Kaaks,<sup>30</sup> Tatsuhiko Kubo,<sup>31</sup> Loic Le Marchand,<sup>32</sup> Robert J MacInnis,<sup>22,23</sup> Satu Männistö,<sup>33</sup> E Jeffrey Metter,<sup>34</sup> Kazuya Mikami,<sup>35</sup> Lorelei A Mucci,<sup>24</sup> Anja W Olsen <sup>36,37</sup>, Kotaro Ozasa <sup>38</sup>, Domenico Palli,<sup>39</sup> Kathryn L Penney,<sup>24,25</sup> Elizabeth A Platz <sup>40</sup>, Michael N Pollak,<sup>41</sup> Monique J Roobol,<sup>42</sup> Catherine A Schaefer,<sup>43</sup> Jeannette M Schenk,<sup>44</sup> Pär Stattin <sup>45</sup>, Akiko Tamakoshi,<sup>46</sup> Elin Thysell,<sup>47</sup> Chiaojung Jillian Tsai,<sup>48</sup> Mathilde Touvier <sup>17</sup>, Stephen K Van Den Eeden,<sup>43,49</sup> Elisabete Weiderpass <sup>50</sup>, Stephanie J Weinstein,<sup>11</sup> Lynne R Wilkens,<sup>32</sup> Bu B Yeap<sup>19,51</sup>; The PRACTICAL Consortium, CRUK, BPC3, CAPS, PEGASUS<sup>‡</sup>, Naomi E Allen<sup>4,52†</sup>, Timothy J Key,<sup>1†</sup> and 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>Genomic Epidemiology Branch, International Agency for Research on Cancer, Lyon, France, <sup>3</sup>Section of Nutrition and Metabolism, International Agency for Research on Cancer, Lyon, France, <sup>4</sup>Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford, UK, <sup>5</sup>Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK, <sup>6</sup>Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK, <sup>7</sup>MRC Integrative Epidemiology Unit (IEU), Bristol Medical School, University of Bristol, Bristol, UK, <sup>8</sup>National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and Weston NHS Foundation Trust and University of Bristol, Bristol, UK, <sup>9</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK, <sup>10</sup>Department of Hygiene and Epidemiology, University of

Ioannina School of Medicine, Ioannina, Greece, <sup>11</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA, <sup>12</sup>Group of Epidemiology of Cancer and Other Chronic Diseases, Navarra Public Health Institute, Pamplona, Spain, <sup>13</sup>Group of Epidemiology of Cancer and Other Chronic Diseases, Navarra Institute for Health Research (IdiSNA), Pamplona, Spain, <sup>14</sup>CIBER Epidemiology and Public Health CIBERESP, Madrid, Spain, <sup>15</sup>Centre for Nutrition, Prevention and Health Services, National Institute for Public Health and the Environment (RIVM), Utrecht, The Netherlands, <sup>16</sup>Child Health and Development Studies, Public Health Institute, Berkeley, CA, USA, <sup>17</sup>Sorbonne Paris Nord University, Nutritional Epidemiology Research Team, Epidemiology and Statistics Research Center, University of Paris, Bobigny, France, <sup>18</sup>National Institute on Aging, Baltimore, MD, USA, <sup>19</sup>WA Centre for Health & Ageing, Medical School, University of Western Australia, Perth, WA, Australia, <sup>20</sup>Western Australian Centre for Health and Ageing, University of Western Australia, Perth, WA, Australia, <sup>21</sup>Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia, <sup>22</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia, <sup>23</sup>Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Melbourne, VIC, Australia, <sup>24</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA, <sup>25</sup>Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, <sup>26</sup>Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA, <sup>27</sup>Department of Preventive Medicine, Center for Genetic Epidemiology, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA, USA, <sup>28</sup>IGFs & Metabolic Endocrinology Group, Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK, <sup>29</sup>National Clinical Research Center for Metabolic Diseases, Key Laboratory of Diabetes Immunology, Ministry of Education, and Department of Metabolism and Endocrinology, Second Xiangya Hospital of Central South University, Changsha, Hunan, China, <sup>30</sup>Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany, <sup>31</sup>Department of Public Health and Health Policy, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan, <sup>32</sup>University of Hawaii, Cancer Center, Honolulu, HI, USA, <sup>33</sup>Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland, <sup>34</sup>Department of Neurology, College of Medicine, University of Tennessee Health Science Center, Memphis, TN, USA, <sup>35</sup>Department of Urology, Japanese Red Cross Kyoto Daiichi Hospital, Kyoto, Japan, <sup>36</sup>Department of Public Health, Aarhus University, Aarhus, Denmark, <sup>37</sup>Danish Cancer Society, Research Center, Copenhagen, Denmark, <sup>38</sup>Department of Epidemiology, Radiation Effects Research Foundation, Hiroshima, Japan, <sup>39</sup>Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network, Florence, Italy, <sup>40</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, <sup>41</sup>Departments of Medicine and Oncology, McGill University, Montreal, QC, Canada, <sup>42</sup>Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands, <sup>43</sup>Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA, <sup>44</sup>Cancer Prevention Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA, <sup>45</sup>Department of Surgical Sciences, Uppsala University, Uppsala, Sweden, <sup>46</sup>Department of Public Health, Faculty of Medicine, Hokkaido University, Sapporo, Japan, <sup>47</sup>Department of Medical Biosciences, Umeå University, Umeå, Sweden, <sup>48</sup>Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA, <sup>49</sup>Department of Urology, University of California San Francisco, San Francisco, CA, USA, <sup>50</sup>Director's Office, International Agency for Research on Cancer, World Health Organization, Lyon, France, <sup>51</sup>Department of Endocrinology and Diabetes, Fiona Stanley Hospital, Perth, WA, Australia and <sup>52</sup>UK Biobank Ltd, Stockport, UK

\*Corresponding author. Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. E-mail: [eleanor.watts@nih.gov](mailto:eleanor.watts@nih.gov)

†These authors contributed equally.

‡Members from the PRACTICAL Consortium, CRUK, BPC3, CAPS and PEGASUS are acknowledged at the end of the article.

## Abstract

**Background:** Previous studies had limited power to assess the associations of circulating insulin-like growth factors (IGFs) and IGF-binding proteins (IGFBPs) with clinically relevant prostate cancer as a primary endpoint, and the association of genetically predicted IGF-I with aggressive prostate cancer is not known. We aimed to investigate the associations of IGF-I, IGF-II, IGFBP-1, IGFBP-2 and IGFBP-3 concentrations with overall, aggressive and early-onset prostate cancer.

**Methods:** Prospective analysis of biomarkers using the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group dataset (up to 20 studies, 17 009 prostate cancer cases, including 2332 aggressive cases). Odds ratios (OR) and 95% confidence intervals (CI) for prostate cancer were estimated using conditional logistic regression. For IGF-I, two-sample Mendelian randomization (MR) analysis was undertaken using instruments identified using UK Biobank (158 444 men) and outcome data from PRACTICAL (up to 85 554 cases, including 15 167 aggressive cases). Additionally, we used colocalization to rule out confounding by linkage disequilibrium.

**Results:** In observational analyses, IGF-I was positively associated with risks of overall (OR per 1 SD = 1.09; 95% CI 1.07, 1.11), aggressive (1.09: 1.03, 1.16) and possibly early-onset disease (1.11: 1.00, 1.24); associations were similar in MR analyses (OR per 1 SD = 1.07: 1.00, 1.15; 1.10: 1.01, 1.20; and 1.13; 0.98, 1.30, respectively). Colocalization also indicated a shared signal for IGF-I and prostate cancer (PP4: 99%). Men with higher IGF-II (1.06: 1.02, 1.11) and IGFBP-3 (1.08: 1.04, 1.11) had higher risks of overall prostate cancer, whereas higher IGFBP-1 was associated with a lower risk (0.95: 0.91, 0.99); these associations were attenuated following adjustment for IGF-I.

**Conclusions:** These findings support the role of IGF-I in the development of prostate cancer, including for aggressive disease.

**Key words:** Insulin-like growth factor-I, prostate cancer, aggressive prostate cancer, prospective analysis, Mendelian randomization, international consortia

### Key Messages

- We used observational and genetic data from international consortia to investigate the associations of circulating insulin-like growth factors (IGF-I, IGF-II) and their binding proteins (IGFBP-1,-2,-3) with overall, aggressive and early-onset prostate cancer.
- Our findings support the role of IGF-I in the development of prostate cancer, including aggressive disease.
- Our results suggest the need for more research on the modifiable determinants of IGF-I, and whether interventions to lower IGF-I might reduce the risk of prostate cancer.

## Introduction

Prostate cancer is the second most common cancer in men worldwide and a leading cause of cancer death.<sup>1</sup> Insulin-like growth factors (IGFs) are important growth-promoting peptides that act through the IGF-I receptor.<sup>2,3</sup> IGF-I and IGF-II are mainly produced by the liver and circulate in the bloodstream, but they are also produced in local tissues where they function in a paracrine/autocrine

manner.<sup>3</sup> The majority of both of these growth factors circulate bound to IGF proteins (IGFBPs),<sup>2,4</sup> which extend the half-life of the IGFs and modulate IGF signalling.<sup>2,4</sup> Higher IGF-I signalling increases cell survival and decreases apoptosis, increasing the probability of carcinogenesis.<sup>4,5</sup> Circulating IGF-I concentrations are positively associated with risks of several cancers, particularly prostate, breast and colorectal cancer.<sup>6,7</sup>

The Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG) is a pooled individual participant nested case-control dataset of prospective studies of hormonal and nutritional factors and prostate cancer risk, which previously reported positive associations of IGF-I, IGF-II, IGFBP-2 and IGFBP-3 with overall prostate cancer risk and an inverse association with IGFBP-1.<sup>8</sup> However, in this previous study it was unclear whether IGF-II or the IGFBPs are associated with prostate cancer independently of IGF-I, and the analyses of associations with aggressive disease subtypes were underpowered to provide strong evidence of an effect.<sup>8</sup> The EHNBPCCG dataset has since been expanded to include more than double the number of prostate cancer cases (up to 17 000 prostate cancer cases, including 2300 aggressive cases).

In blood-based observational analyses it is difficult to rule out the possibility of biases including residual confounding or reverse causality. Mendelian randomization (MR) uses germline genetic variants as proxies of putative risk factors and estimates their associations with disease risk. These germline genetic variants are randomly allocated and fixed at conception, and therefore MR is less likely to be affected by these biases and so is potentially a more robust method for causal inference.<sup>9</sup> In order to appraise causality for IGF-I, we carried out two-sample MR analyses using instruments identified from UK Biobank and genetic data from the PRACTICAL consortium.<sup>10–12</sup> Using these genetic datasets, we also ran colocalization analyses to investigate whether the *IGF1* gene region and prostate cancer share the same genetic signal to exclude the possibility of confounding by linkage disequilibrium.<sup>13</sup>

Using these two international consortia and UK Biobank, we aimed to assess the associations of circulating IGF-I with overall, aggressive and early-onset prostate cancer risk, using observational and genetic methods. The analysis of very large datasets can provide more robust risk estimates, and the integration of evidence from these different epidemiological approaches can strengthen the basis for causal inference.<sup>14</sup> We additionally report observational associations of IGF-II and IGFBPs-1,-2,-3 with overall, aggressive and early-onset subtypes.

## Methods

### Endogenous hormones, nutritional biomarkers and Prostate Cancer Collaborative Group

#### Data collection and study designs

Individual participant data were available from up to 20 prospective studies with IGF-I (17 009 cases), IGF-II (4466 cases), IGFBP-1 (4491 cases), IGFBP-2 (3776 cases)

and IGFBP-3 (9113 cases) measurements. Participating studies are listed in [Supplementary Table S1](#) and further details of data collection and processing are provided in the [Supplementary material](#). Matching criteria are shown in [Supplementary Table S2](#). Assay details and hormone measurement data are provided in [Supplementary Table S3](#).

#### Data processing and outcomes

Disease definitions were as defined by the PRACTICAL consortium.<sup>10,11</sup> Aggressive prostate cancer was categorized as ‘yes’ for any of the following: disease metastases at diagnosis (M1), Gleason score 8+ (or equivalent), prostate cancer death (defined as death from prostate cancer) or prostate-specific antigen (PSA) >100 ng/mL. Early-onset prostate cancer was defined as a diagnosis aged ≤55 years. Further details of the disease characterization can be found in the [Supplementary Methods](#).

#### Statistical analysis

Conditional logistic regression was used to estimate prostate cancer risk by circulating concentrations of IGF-I, IGF-II, IGFBP-1, IGFBP-2 and IGFBP-3. Analyses were conditioned on the study-specific matching variables and adjusted for age at blood collection, body mass index (BMI), height, smoking status, alcohol consumption, racial or ethnic group, education, married/cohabiting and diabetes status. Biomarkers were standardized by study and entered into the model as continuous variables, so each increment represents 1 study-specific SD increase in biomarker concentration. For categorical analyses, biomarkers were categorized into study-specific fifths with cut-points determined in controls.<sup>15</sup> Further details are available in the [Supplementary Methods](#).

#### Further analyses

We examined heterogeneity in the associations of each biomarker with prostate cancer by participant characteristics, with subgroups defined a priori based on the availability of data and previous analyses using this dataset<sup>8,16</sup>; heterogeneity in the associations by study was also examined ([Supplementary Methods](#)). We additionally investigated unadjusted matched associations, associations in tenths, and estimates per 80th percentile increase. Associations were also examined following mutual adjustment for other biomarkers (IGF-I, IGF-II, IGFBP-1,-2,-3, free and total testosterone and sex hormone-binding globulin [SHBG]), and we tested for interactions between these biomarkers; further details are available in the [Supplementary Methods](#). Stratified analyses and associations in tenths were not investigated for early-onset disease due to the limited number of cases.

## Mendelian randomization analysis

### Genetic instruments for hormone concentrations

Single nucleotide polymorphisms (SNPs) associated with circulating IGF-I concentrations were identified from a publicly available genome-wide association study (GWAS) based on 158 444 male UK Biobank participants of White British ancestry ( $P < 5 \times 10^{-8}$  significance threshold).<sup>17</sup> We pruned SNPs by a linkage disequilibrium threshold of  $r^2 < 0.001$ , based on the lowest  $P$ -value.

### Genetic associations with prostate cancer

SNP associations for prostate cancer were obtained from the PRACTICAL and GAME-ON/ELLIPSE consortia,<sup>10,11</sup> which currently do not include UK Biobank data. Individual studies included in these consortia are detailed in Conti *et al.*<sup>12</sup> and Schumacher *et al.*<sup>10</sup> Associations with overall prostate cancer risk were generated from 85 554 prostate cancer cases and 91 972 controls,<sup>12</sup> with aggressive disease from 15 167 cases and 58 308 controls and with early-onset disease from 6988 cases and 44 256 controls,<sup>10</sup> all of White European ancestry.

### Statistical analysis

The MR estimation for hormones was conducted using the inverse-variance weighted (IVW) method.<sup>18</sup> We additionally calculated the  $I^2$  statistic to assess measurement error in SNP-exposure associations,<sup>19</sup> the  $F$  statistic to assess instrument strength,<sup>20,21</sup> Cochran's  $Q$  statistic to test for heterogeneity between the MR estimates for each SNP<sup>22</sup> and PhenoScanner was used to assess pleiotropy of the genetic instruments.<sup>23</sup> As sensitivity analyses, we used the MR residual sum and outlier (MR-PRESSO), MR robust adjusted profile score (MR-RAPS) and leave-one-out analyses to investigate the role of SNP outliers.<sup>24</sup> To assess pleiotropy, we used the weighted median, MR-Egger and the MR-Egger intercept.<sup>25</sup> We also used the contamination mixture method, which assumes a normal distribution of valid instruments around the true causal value, and invalid instruments are normally distributed around zero in order to account for potentially pleiotropic variants.<sup>26</sup> To rule out reverse causality, analyses were repeated after applying Steiger filtering which excludes variants with larger effects on prostate cancer risk than on IGF-I.<sup>27</sup>

The associations of the IGF-I *cis*-SNP, defined as the lead SNP on the biomarker gene coding region identified from the exposure datasets, with prostate cancer were investigated using the Wald ratio. This *cis*-SNP is less likely than *trans*-SNPs to be affected by horizontal pleiotropy.<sup>28</sup>

## Colocalization analysis

Colocalization was used to investigate whether the associations of variation in the *IGF1* gene region with both circulating IGF-I concentration and prostate cancer risk, share the same genetic signal or whether the associations identified by our MR analysis may be confounded by linkage disequilibrium.<sup>13</sup> Analyses were conducted for a 75-kb region surrounding the lead IGF-I *cis*-SNP (rs5742653) using the UK Biobank and PRACTICAL datasets.<sup>12,17</sup> Colocalization was assessed using three approaches: conventional colocalization,<sup>13</sup> which tests for the presence of a single shared genetic signal; as well as the sum of single effects (SuSiE) regression framework<sup>29</sup>; and conditional iterative colocalization.<sup>30</sup> The latter two methods allow for the possibility of multiple independent (but partially correlated) causal variants in proximity.<sup>31</sup> We created colocalization plots using LocusCompareR<sup>32</sup> and a z-z locus plot.<sup>33</sup> We considered a posterior probability of a shared causal variant (PP4) of  $> 0.7$  as being consistent with evidence of colocalization between IGF-I and prostate cancer.<sup>13</sup> Further details of the colocalization analysis are available in the [Supplementary Methods](#).

Details of statistical software and packages used are available in the [Supplementary Methods](#). All tests of significance were two-sided, and  $P$ -values  $< 0.05$  were considered statistically significant.

## Results

### Study and participant characteristics in the observational analyses

A total of 20 studies, contributing up to 17 009 cases and 37 243 controls, were included in this analysis. Prostate cancer was classified as aggressive in 2332 cases and early-onset disease in 607 cases. Study participants were 91.3% of White ethnicity ([Table 1](#)). Men who were diagnosed with overall prostate cancer were taller and had a lower BMI than their matched controls ([Table 1](#)).

Prostate cancer characteristics by study are displayed in [Supplementary Table S4](#). Mean age at blood collection for each study ranged from 33.8 to 76.8 years (overall mean = 61.2 years, SD = 7.8 years). Cases were diagnosed on average 6.7 years (SD = 5.4) after blood collection, and the average age at diagnosis was 67.5 years (SD = 6.5) ([Table 1](#)). Aggressive disease was diagnosed on average 8.0 years after blood collection (SD = 6.3) ([Table 1](#)). Partial correlations between biomarkers ranged from  $r = -0.004$  (PSA and IGF-II) to  $r = 0.54$  (IGF-II and IGFBP-2) ([Supplementary Table S5](#)).

**Table 1** Characteristics of prostate cancer cases and controls in the EHNBPCCG participants

	Controls	Cases		
		Overall	Aggressive <sup>a</sup>	Early-onset <sup>b</sup>
N	37 243	17 009	2332	607
Age (years), mean (SD)	61.4 (7.7)	60.7 (8.0)	61.2 (7.9)	47.1 (5.3)
Height (cm), mean (SD)	174.9 (7.0)	175.3 (7.1)	175.2 (7.3)	177.3 (6.9)
BMI (kg/m <sup>2</sup> ), mean (SD)	27.4 (4.1)	26.8 (3.6)	26.9 (3.9)	26.3 (3.6)
PSA at blood collection (ng/mL), mean (IQR)	0.9 (1.2)	2.4 (3.3)	2.9 (5.7)	1.9 (2.8)
Time from blood collection to diagnosis, mean (SD)	–	6.7 (5.4)	8.0 (6.3)	5.6 (5.0)
Age at diagnosis, mean (SD)	–	67.5 (6.5)	67.3 (6.2)	52.7 (2.4)
Racial/ethnic group, N (%)				
White	33 988 (91.3)	15 617 (91.8)	2217 (95.1)	532 (87.6)
Black	1145 (3.1)	505 (3.0)	53 (2.3)	30 (4.9)
East Asian	336 (0.9)	146 (0.9)	8 (0.3)	3 (0.5)
Other	707 (1.9)	266 (1.6)	22 (0.9)	11 (1.8)
Not known	1067 (2.9)	475 (2.8)	32 (1.4)	31 (5.1)
Smoking status, N (%)				
Never	14 985 (40.2)	6791 (39.9)	804 (34.5)	305 (50.2)
Ex	16 511 (44.3)	7300 (42.9)	1000 (42.9)	170 (28.0)
Current	5203 (14.0)	2533 (14.9)	491 (21.1)	127 (20.9)
Not known	544 (1.5)	385 (2.3)	37 (1.6)	5 (0.8)
Alcohol consumption (g ethanol/day), N (%)				
Non-drinker	2851 (7.7)	1806 (10.6)	264 (11.3)	52 (8.6)
<10	9073 (24.4)	4535 (26.7)	649 (27.8)	162 (26.7)
10 +	21 385 (57.4)	9171 (53.9)	1284 (55.1)	346 (57.0)
Not known	3934 (10.6)	1497 (8.8)	135 (5.8)	47 (7.7)
Diabetes status, N (%)				
Yes	2921 (7.8)	864 (5.1)	127 (5.4)	12 (2.0)
No	31 707 (85.1)	14 847 (87.3)	2052 (88.0)	533 (87.8)
Not known	2615 (7.0)	1298 (7.6)	153 (6.6)	62 (10.2)
Married/cohabiting, N (%)				
Yes	9478 (25.4)	6810 (40.0)	1157 (49.6)	235 (38.7)
No	1407 (3.8)	922 (5.4)	149 (6.4)	40 (6.6)
Not known	26 358 (70.8)	9277 (54.5)	1026 (44.0)	332 (54.7)

BMI, body mass index; EHNBPCCG, Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group; IQR, interquartile range; PSA, prostate-specific antigen; SD, standard deviation.

<sup>a</sup>Aggressive disease was defined as Gleason Score 8+, death from prostate cancer, metastatic disease or PSA >100 ng/mL.

<sup>b</sup>Onset defined as diagnosed aged ≤55 years.

## IGF-I

In observational analyses, higher IGF-I was related to dose-dependent elevated risks of overall (OR per 1 SD increment = 1.09; 95% CI 1.07, 1.11;  $P < 0.0001$ ) and aggressive prostate cancer (1.09: 1.03, 1.16;  $P = 0.01$ ), and there was a suggestive association with early-onset disease (1.11: 1.00, 1.24;  $P = 0.05$ ).

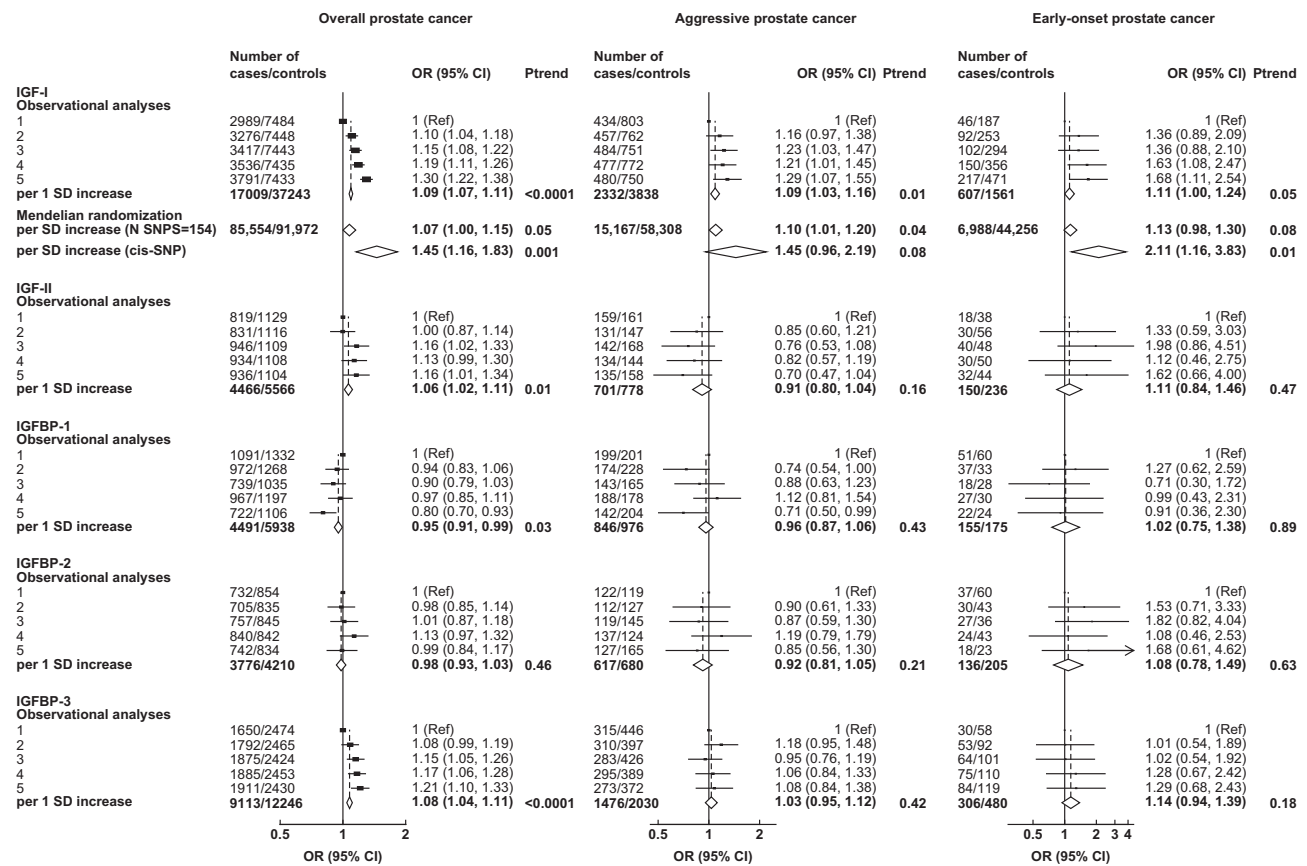
In MR analyses, higher IGF-I was associated with increased risks of overall and aggressive disease (OR per genetically predicted 1-SD increment = 1.07: 1.00, 1.15;  $P = 0.05$ ; and 1.10: 1.01, 1.20;  $P = 0.04$ , respectively) and was positively related to risk of early-onset disease (1.13: 0.98, 1.30;  $P = 0.08$ ) (Figure 1). The MR sensitivity analyses were generally directionally consistent with IGF-I, although the confidence intervals were wider (Table 2).

The associations with prostate cancer risk were also directionally consistent when IGF-I was proxied by the *cis*-SNP (rs5742653) (1.45: 1.16, 1.83;  $P = 0.001$ ; 1.45: 0.96–2.19;  $P = 0.08$ ; and 2.11: 1.16, 3.83;  $P = 0.01$ , for overall, aggressive and early-onset disease, respectively) (Figure 1). Both SuSiE and conditional iterative analyses indicated multiple independent shared causal variants for IGF-I and overall prostate cancer (maximum PP4 >0.99 using SuSiE and PP4 = 0.72 using conditional iterative regression) (Supplementary Table S6 and Supplementary Figures S1 and S2).

## IGF-II and IGFBPs-1–3

In observational analyses, men with higher circulating IGF-II and IGFBP-3 had an elevated risk of overall prostate





**Figure 1** Risks of overall, aggressive\* and early-onset† prostate cancer by study-specific fifths of biomarker concentrations (observational only) and 1 SD increment (observational and Mendelian randomization). Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, education status, racial/ethnic group and diabetes status. The position of each square indicates the magnitude of the odds ratio, and the area of the square is proportional to the inverse of the variance of the log odds ratio. The length of the horizontal line through the square indicates the 95% confidence interval. MR risk estimates are estimated using the inverse-variance weighted method for the full instrument methods and the Wald ratio in the *cis*-SNP analyses. *P*<sub>trend</sub> represents 1-SD increase in biomarker concentration. \*Aggressive cancer defined as Gleason grade 8+, or prostate cancer death, or metastases or PSA >100 ng/mL. †Early-onset defined as diagnosed ≤55 years. BMI, body mass index; CI, confidence interval; IGF, insulin-like growth factor; IGFBP, insulin-like growth factor-binding protein; OR, odds ratio; PSA, prostate-specific antigen; SD, standard deviation; MR, Mendelian randomization; SNP, single nucleotide polymorphism

cancer (OR per 1-SD increment = 1.06; 95% CI 1.02, 1.11; *P* = 0.01; and 1.08: 1.04, 1.11; *P* < 0.0001, respectively). IGFBP-1 was inversely associated with overall prostate cancer (0.95: 0.91, 0.99; *P* = 0.03), and IGFBP-2 was not associated with prostate cancer risk (0.98: 0.93, 1.03; *P* = 0.46) (Figure 1). These biomarkers were not associated with aggressive or early-onset disease (Figure 1).

**Further analyses—observational analysis**

Associations of IGF-I with overall and aggressive prostate cancer were generally consistent by subgroups and secondary outcomes (Figures 2 and 3). The OR for prostate cancer death was 1.08 for IGF-I (1.00, 1.17) (Figure 2). There was some evidence of larger magnitudes of associations with overall prostate cancer for men with a family history of prostate cancer (1.19: 1.09, 1.29) than for men without (1.07: 1.03, 1.11; *P*<sub>het</sub> = 0.02) (Figure 2).

The associations of IGF-II and IGFBPs with prostate cancer risk were broadly similar by subgroups (Supplementary Figures S3–S10). There was evidence of heterogeneity in the association of IGFBP-2 with overall prostate cancer by BMI (*P*<sub>het</sub> = 0.0007); for men whose BMI was <25 kg/m<sup>2</sup> at baseline, IGFBP-2 was inversely associated with prostate cancer (0.89: 0.83, 0.96), and the OR for men with BMI 30+ was 1.19 (0.99, 1.42) (Supplementary Figure S7). IGFBP-2 was also inversely associated with aggressive disease risk for men whose BMI was <25 kg/m<sup>2</sup> (0.78: 0.66, 0.94), but not for men who had a higher BMI (*P*<sub>het</sub> = 0.01) (Supplementary Figure S8).

Associations with overall and aggressive prostate cancer by study are available in Supplementary Figures S11–S20. There was some evidence of heterogeneity by study in the associations of IGF-I with aggressive disease (*P*<sub>het</sub> = 0.02) (Supplementary Figure S12), and IGF-II and IGFBP-2 with overall prostate cancer risk (*P*<sub>het</sub> = 0.0001 and 0.02,

**Table 2** Mendelian randomization estimates between genetically predicted circulating IGF-I concentrations and overall, aggressive and early-onset prostate cancer

	Variance explained	N SNPs	Overall prostate cancer (85 554 cases, 91 972 controls)		Aggressive prostate cancer <sup>a</sup> (15 167 cases, 58 308 controls)		Early-onset prostate cancer <sup>b</sup> (6988 cases, 44 256 controls)	
			OR per 1-SD increment (95% CI)	P-value	OR per 1-SD increment (95% CI)	P-value	OR per 1-SD increment (95% CI)	P-value
IGF-I (SD = 5.4 nmol/L)								
Inverse-variance weighted	8.7%	154	1.07 (1.00, 1.15)	0.05	1.10 (1.01, 1.20)	0.04	1.13 (0.98, 1.30)	0.08
Weighted median			1.01 (0.95, 1.08)	0.71	1.03 (0.91, 1.16)	0.63	1.07 (0.90, 1.29)	0.44
MR-Egger			1.00 (0.85, 1.17)	0.99	1.01 (0.83, 1.24)	0.90	0.98 (0.71, 1.35)	0.89
MR-Egger intercept				0.73		0.38		0.31
MR-RAPS			1.04 (0.98, 1.12)	0.22	1.11 (1.00, 1.22)	0.04	1.11 (0.96, 1.28)	0.16
MR-PRESSO			1.06 (1.00, 1.12)	0.05	1.08 (0.99, 1.18)	0.08	1.10 (0.97, 1.25)	0.13
Contamination mixture			1.01 (0.90, 1.06)	0.73	1.32 (1.17, 1.45)	0.0005	1.13 (0.96, 1.42)	0.16
<i>cis</i> -SNP (rs5742653)	0.2%	1	1.45 (1.16, 1.83)	0.001	1.45 (0.96, 2.19)	0.08	2.11 (1.16, 3.83)	0.01

SD estimates based on UK Biobank males.

CI, confidence interval; IGF-I, insulin-like growth factor-I; MR, Mendelian randomization; OR, odds ratio; PRESSO, pleiotropy residual sum and outlier; PSA, prostate-specific antigen; RAPS, robust adjusted profile score; SD, standard deviation; SNP, single nucleotide polymorphism.

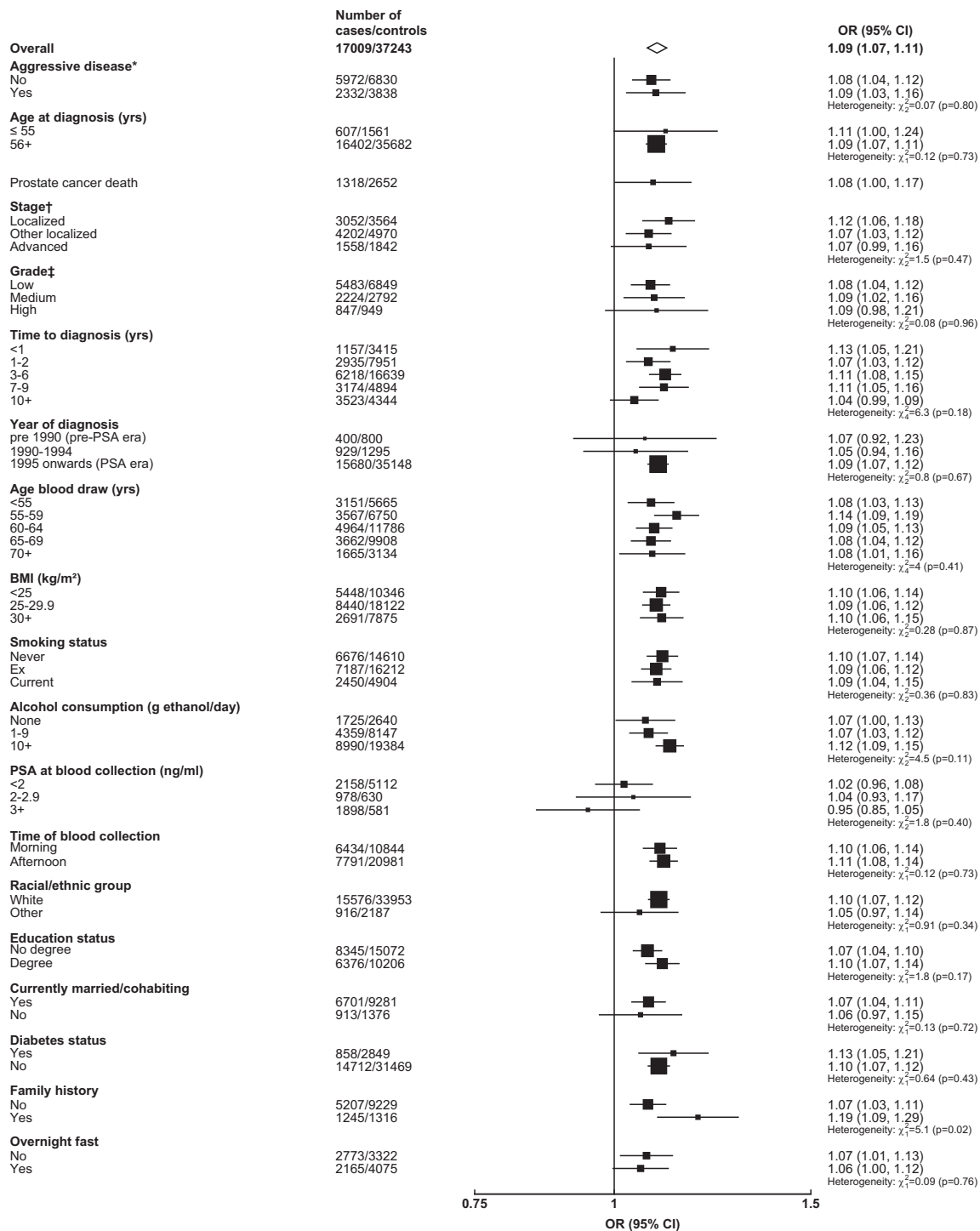
<sup>a</sup>Aggressive disease was defined as Gleason Score 8+, death from prostate cancer, metastatic disease or PSA > 100 ng/mL.<sup>b</sup>Early-onset defined as diagnosed aged ≤ 55 years.

respectively) (Supplementary Figures S13 and S17). Associations were broadly similar to the primary analyses in unadjusted matched analyses (Supplementary Figure S21), using study-specific tenths (Supplementary Figure S22) and per 80 percentile increase (Supplementary Table S7). Following mutual adjustment for IGF-I, the associations of IGF-II and IGFBP-1 with risk were attenuated to the null (Supplementary Table S8). For IGF-I and IGFBP-3, mutual adjustment slightly attenuated the associations with overall prostate cancer risk, but both these associations remained (Supplementary Table S8).

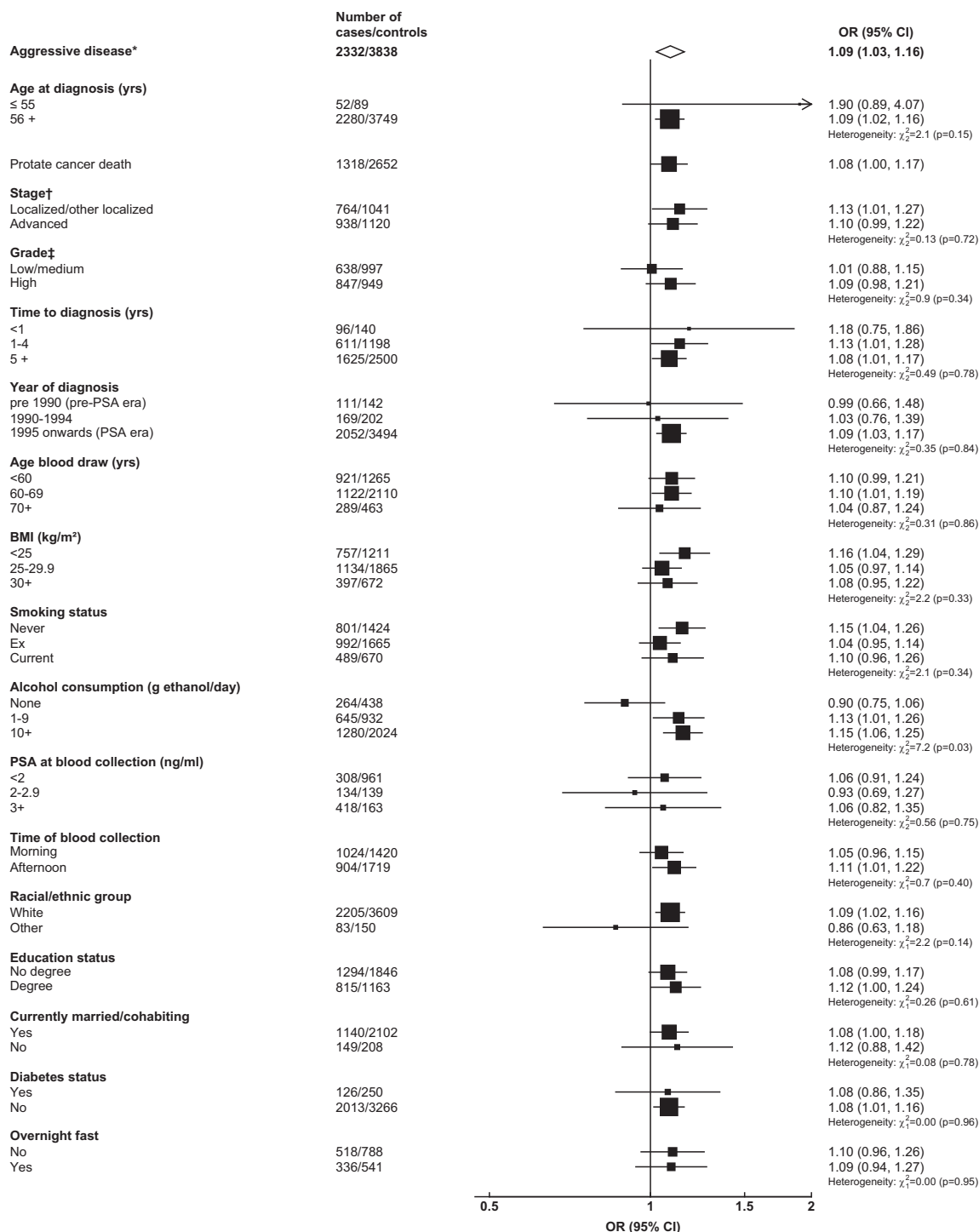
There was some evidence of interactions in the associations of IGF-II, IGFBP-1 and IGFBP-2 concentrations with prostate cancer risk by total testosterone concentrations; men with total testosterone concentrations above the study-specific median showed evidence of a positive relationship for IGF-II and an inverse association for IGFBP-1, whereas these associations were null for men with lower total testosterone concentrations ( $P_{\text{het}} = 0.03$  and  $0.02$ , respectively) (Supplementary Table S9). Only men with lower total testosterone concentrations had a positive association between IGFBP-2 and overall prostate cancer ( $P_{\text{het}} = 0.01$ ). For aggressive disease, the OR for IGFBP-2 was 1.27 for men with lower total testosterone concentrations (1.00, 1.62), and in men with higher total testosterone there was an inverse relationship of IGFBP-2 with aggressive disease (0.75: 0.60, 0.93;  $P_{\text{het}} < 0.01$ ), although the number of aggressive cases was limited ( $N = 443$ ) (Supplementary Table S10).

#### Further analyses—mendelian randomization

There was no strong evidence of measurement error in the genetic instruments for IGF-I ( $I^2 = 0.99$ ) and all SNPs had an F statistic > 10.<sup>20</sup> There was significant heterogeneity in the MR estimates for the SNPs with overall prostate cancer, and for aggressive and early-onset disease (Cochran's Q  $P < 0.001$ ). Full MR results are found in Supplementary Table S11. Forest plots of single-SNP results are available in Supplementary Figures S23–25, leave-one-out plots are available in Supplementary Figures S26–28 and MR scatterplots are available in Supplementary Figure S29. Outliers identified by MR-PRESSO are available in Supplementary Table S12. Following Steiger filtering, the results were slightly attenuated (Supplementary Table S13). Using PhenoScanner, 430 traits were identified as being linked to genetically predicted IGF-I, including height and measures of adiposity (Supplementary Figure S30). Higher concentrations of IGF-I instrumented by the *cis*-SNP (rs5742653) were associated with increased peak expiratory flow ( $P < 5 \times 10^{-8}$ ).



**Figure 2** Odds ratio (95% CIs) for overall prostate cancer per study-specific 1-SD increment of IGF-I concentration by subgroup in the EHNBPCCG. Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, education status, racial/ethnic group and diabetes status. The position of each square indicates the magnitude of the odds ratio, and the area of the square is proportional to the inverse of the variance of the log odds ratio. The length of the horizontal line through the square indicates the 95% confidence interval. Tests for heterogeneity for case-defined factors were obtained by fitting separate models for each subgroup and assuming independence of the ORs using a method analogous to a meta-analysis. Tests for heterogeneity for non-case-defined factors were assessed with a  $\chi^2$  test of interaction between subgroup and the binary variable. \*Aggressive cancer defined as Gleason grade 8+, or prostate cancer death, or metastases or PSA >100 ng/mL. †Localized defined as TNM stage <T2 with no reported lymph node involvement or metastases or stage I; other localized stage if TNM stage T2 with no reported lymph node involvement or metastases, stage II, or equivalent; advanced stage if they were TNM stage T3 or T4 and/or N1+ and/or M1, stage III-IV or equivalent. ‡ Low grade defined as Gleason score was <7 or equivalent (i.e. extent of differentiation good, moderate); medium grade if Gleason score was 7 (i.e. poorly differentiated); high grade if the Gleason score was  $\geq 8$  or equivalent (i.e. undifferentiated). BMI, body mass index; CI, confidence interval; EHNBPCCG, Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group; IGF-I, insulin-like growth factor-I; OR, odds ratio; PSA, prostate-specific antigen; SD, standard deviation; TNM, tumour, node, metastases



**Figure 3** Odds ratio (95% CIs) for aggressive\* prostate cancer per study-specific 1-SD increment of IGF-I concentration by subgroup in the EHNBPCCG. Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, education status, racial/ethnic group and diabetes status. The position of each square indicates the magnitude of the odds ratio, and the area of the square is proportional to the inverse of the variance of the log odds ratio. The length of the horizontal line through the square indicates the 95% confidence interval. Tests for heterogeneity for case-defined factors were obtained by fitting separate models for each subgroup and assuming independence of the ORs using a method analogous to a meta-analysis. Tests for heterogeneity for non-case-defined factors were assessed with a  $\chi^2$  test of interaction between subgroup and the binary variable. \*Aggressive cancer defined as Gleason grade 8+, or prostate cancer death, or metastases or PSA > 100 ng/mL. †Localized/other localized defined as TNM stage <T2 with no reported lymph node involvement or metastases or stage I, or TNM stage T2 with no reported lymph node involvement or metastases, stage II, or equivalent; advanced stage if they were TNM stage T3 or T4 and/or N1+ and/or M1, stage III-IV or equivalent. ‡Low/medium grade defined as Gleason score was <8 or equivalent (i.e. extent of differentiation good, moderate, poor); high grade if the Gleason score was  $\geq 8$  or equivalent (i.e. undifferentiated). BMI, body mass index; CI, confidence interval; EHNBPCCG, Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group; IGF, insulin-like growth factor-I; OR, odds ratio; PSA, prostate-specific antigen; SD, standard deviation; TNM, tumour, node, metastases

## Discussion

This is the first study that has applied both observational and genetic approaches using data from large international consortia to investigate the associations of IGF-I with prostate cancer risk. Our results support a role of circulating IGF-I in the development of prostate cancer, including aggressive disease. In observational analyses, IGF-II and IGFFBPs-1 and -3 were also associated with overall prostate cancer risk, but these associations were attenuated following adjustment for IGF-I.

Genetic analyses may be more informative than observational analyses about the direct role of the exposure on the outcome. The weaker findings from genetic analyses from the multi-SNP (*cis* and *trans*) instrument, compared with the *cis*-SNP may be related to associations of some of the *trans*-SNPs with other components of the IGF signalling pathway such as the IGFFBPs.<sup>34</sup> For the lead *cis*-SNP MR we observed larger magnitude effects, which likely indicates stronger biological plausibility of a direct role for IGF-I and a reduced role of horizontal pleiotropy,<sup>35</sup> and may also be due to the possible role of local *IGF1* expression in the prostate tissue. Moreover, colocalization analyses showed strong evidence of a shared genetic cause at the *IGF1* gene for IGF-I concentrations and risk for prostate cancer, indicating that our findings are unlikely to be due to confounding by linkage disequilibrium.

In our observational analyses, IGF-II, IGFBP-1 and IGFBP-3 were positively associated with overall prostate cancer, but we were underpowered to detect associations with aggressive or early-onset disease. Following further adjustment for IGF-I, the associations with overall disease were attenuated although IGFBP-3 remained significantly associated with overall prostate cancer. These results suggest that the observed associations may be at least partially due to the correlations of these biomarkers with IGF-I. Analogous genetic approaches such as multivariable MR may be useful in exploring the direct and indirect effects of these biomarkers on prostate cancer risk.<sup>36</sup>

These analyses have several strengths. This is the largest collection of observational and genetic data on hormones and prostate cancer risk available, representing almost all the available data worldwide. This large sample size maximizes power to assess associations robustly and enabled us to investigate associations across subgroups. Further, by incorporating observational and genetic methods, we were able to use different lines of evidence for a more robust investigation towards causal inference.<sup>14</sup>

This study had a number of limitations. IGFs and IGFFBPs are also produced locally as well as by the liver, which may affect prostate cancer risk independently of circulating concentrations.<sup>2,4</sup> Consequently, the predictive

value of circulating IGF-I as an indicator of intra-prostatic IGF signalling remains incompletely understood,<sup>4</sup> and future research including measured intra-prostatic IGF-I and IGF-I receptor expression may help to clarify this. Our analyses relied on single biomarker measurements, and although these biomarkers have good reproducibility over a 1 to 5 year period (intraclass correlation coefficients 0.60–0.90 for IGF-I and IGFBP-1,-2,-3),<sup>37–39</sup> this would be expected to lead to underestimates of risk in the observational analyses.<sup>40</sup> Although associations were generally consistent by subgroup, the number of statistical tests in these analyses increased the possibility of false-positives. Assay methods used to measure the biomarkers varied by study, and some IGF biomarkers are more difficult to measure than others (for example, IGF-II); measurement error would be expected to be non-differential and therefore tend to bias associations towards the null. As in the standard approach for MR, effect estimates 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 associations; the concordance of the directions of the associations is therefore particularly important. Wider confidence intervals in MR sensitivity analyses may relate to lower power for some of these methods.<sup>41</sup>

## Conclusion

In conclusion, the findings from these analyses using observational and genetic data from large-scale international consortia are supportive of a role of IGF-I in the aetiology of prostate cancer. For the first time we show evidence that IGF-I is important for aggressive, clinically relevant disease. These findings support the need for more research on the modifiable determinants of IGF-I, and on whether interventions to lower IGF-I might reduce the risk of prostate cancer.

## PRACTICAL, CRUK, BPC3, CAPS and PEGASUS consortia investigators

Principal Investigators from the PRACTICAL [<http://practical.icr.ac.uk/>], CRUK, BPC3, CAPS, PEGASUS consortia: Rosalind A Eeles,<sup>1,2</sup> Christopher A Haiman,<sup>3</sup> Zsofia Kote-Jarai,<sup>1</sup> Fredrick R Schumacher,<sup>4,5</sup> Sara Benlloch,<sup>1,6</sup> Ali Amin Al Olama,<sup>6,7</sup> Kenneth R Muir,<sup>8</sup> Sonja I Berndt,<sup>9</sup> David V Conti,<sup>3</sup> Fredrik Wiklund,<sup>10</sup> Stephen Chanock,<sup>9</sup> Ying Wang,<sup>11</sup> Catherine M Tangen,<sup>12</sup> Jyotsna Batra,<sup>13,14</sup> Judith A Clements,<sup>13,14</sup> APCB BioResource (Australian Prostate Cancer BioResource),<sup>15,14</sup> Henrik Grönberg,<sup>10</sup> Nora Pashayan,<sup>16,17</sup> Johanna Schleutker,<sup>18,19</sup> Demetrius Albanes,<sup>9</sup> Stephanie Weinstein,<sup>9</sup> Alicja Wolk,<sup>20</sup> Catharine

M L West,<sup>21</sup> Lorelei A Mucci,<sup>22</sup> Géraldine Cancel-Tassin,<sup>23,24</sup> Stella Koutros,<sup>9</sup> Karina Dalsgaard Sørensen,<sup>25,26</sup> Eli Marie Grindedal,<sup>27</sup> David E Neal,<sup>28,29,30</sup> Freddie C Hamdy,<sup>31,32</sup> Jenny L Donovan,<sup>33</sup> Ruth C Travis,<sup>34</sup> Robert J Hamilton,<sup>35,36</sup> Sue Ann Ingles,<sup>37</sup> Barry S Rosenstein,<sup>38</sup> Yong-Jie Lu,<sup>39</sup> Graham G Giles,<sup>40,41,42</sup> Robert J MacInnis,<sup>40,41</sup> Adam S Kibel,<sup>43</sup> Ana Vega,<sup>44,45,46</sup> Manolis Kogevinas,<sup>47,48,49,50</sup> Kathryn L Penney,<sup>51</sup> Jong Y Park,<sup>52</sup> Janet L Stanford,<sup>53,54</sup> Cezary Cybulski,<sup>55</sup> Børge G Nordestgaard,<sup>56,57</sup> Sune F Nielsen,<sup>56,57</sup> Hermann Brenner,<sup>58,59,60</sup> Christiane Maier,<sup>61</sup> Jeri Kim,<sup>62</sup> Esther M John,<sup>63</sup> Manuel R Teixeira,<sup>64,65,66</sup> Susan L Neuhausen,<sup>67</sup> Kim De Ruycck,<sup>68</sup> Azad Razack,<sup>69</sup> Lisa F Newcomb,<sup>53,70</sup> Davor Lessel,<sup>71</sup> Radka Kaneva,<sup>72</sup> Nawaid Usmani,<sup>73,74</sup> Frank Claessens,<sup>75</sup> Paul A Townsend,<sup>76,77</sup> Jose Esteban Castelao,<sup>78</sup> Monique J Roobol,<sup>79</sup> Florence Menegaux,<sup>80</sup> Kay-Tee Khaw,<sup>81</sup> Lisa Cannon-Albright,<sup>82,83</sup> Hardev Pandha,<sup>77</sup> Stephen N Thibodeau,<sup>84</sup> David J Hunter,<sup>85</sup> Peter Kraft,<sup>86</sup> William J Blot<sup>87,88</sup> and Elio Riboli<sup>89</sup>

<sup>1</sup>Institute of Cancer Research, London, UK

<sup>2</sup>Royal Marsden NHS Foundation Trust, London, UK

<sup>3</sup>Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA, USA

<sup>4</sup>Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

<sup>5</sup>Seidman Cancer Center, University Hospitals, Cleveland, OH, USA

<sup>6</sup>Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

<sup>7</sup>University of Cambridge, Department of Clinical Neurosciences, Stroke Research Group, Cambridge, UK

<sup>8</sup>Division of Population Health, Health Services Research and Primary Care, University of Manchester, Manchester, UK

<sup>9</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

<sup>10</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

<sup>11</sup>Department of Population Science, American Cancer Society, Atlanta, GA, USA

<sup>12</sup>SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

<sup>13</sup>Australian Prostate Cancer Research Centre-QLD, Institute of Health and Biomedical Innovation and School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD, Australia

<sup>14</sup>Translational Research Institute, Brisbane, QLD, Australia

<sup>15</sup>Australian Prostate Cancer Research Centre-QLD, Queensland University of Technology, Brisbane; Prostate Cancer Research Program, Monash University, Melbourne; Dame Roma Mitchell Cancer Centre, University of Adelaide, Adelaide; Chris O'Brien Lifehouse, Royal Prince Alfred Hospital, Camperdown: Australia

<sup>16</sup>Department of Applied Health Research, University College London, London, UK

<sup>17</sup>Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK

<sup>18</sup>Institute of Biomedicine, University of Turku, Finland

<sup>19</sup>Department of Medical Genetics, Turku University Hospital, Turku, Finland

<sup>20</sup>Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

<sup>21</sup>Division of Cancer Sciences, University of Manchester, Manchester Academic Health Science Centre, Radiotherapy Related Research, Christie Hospital NHS Foundation Trust, Manchester, UK

<sup>22</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>23</sup>CeRePP, Tenon Hospital, Paris, France

<sup>24</sup>Sorbonne Université, GRC n°5, AP-HP, Tenon Hospital, Paris, France

<sup>25</sup>Department of Molecular Medicine, Aarhus University Hospital, Aarhus N, Denmark

<sup>26</sup>Department of Clinical Medicine, Aarhus University, Aarhus N, Denmark

<sup>27</sup>Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

<sup>28</sup>Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford, UK

<sup>29</sup>University of Cambridge, Department of Oncology, Addenbrooke's Hospital, Cambridge, UK

<sup>30</sup>Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Cambridge, UK

<sup>31</sup>Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

<sup>32</sup>Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, UK

<sup>33</sup>Population Health Sciences, Bristol Medical School, University of Bristol, UK

<sup>34</sup>Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>35</sup>Dept. of Surgical Oncology, Princess Margaret Cancer Centre, Toronto, ON, Canada

<sup>36</sup>Dept. of Surgery (Urology), University of Toronto, Toronto, ON, Canada

<sup>37</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA, USA

<sup>38</sup>Department of Radiation Oncology and Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

<sup>39</sup>Centre for Cancer Biomarker and Biotherapeutics, Barts Cancer Institute, Queen Mary University of London, John Vane Science Centre, London, UK

<sup>40</sup>Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia

<sup>41</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia

<sup>42</sup>Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia

<sup>43</sup>Division of Urologic Surgery, Brigham and Womens Hospital, Boston, MA, USA

<sup>44</sup>Fundación Pública Galega Medicina Xenómica, Santiago de Compostela, Spain

<sup>45</sup>Instituto de Investigación Sanitaria de Santiago de Compostela, Santiago De Compostela, Spain

<sup>46</sup>Centro de Investigación en Red de Enfermedades Raras, Madrid, Spain

<sup>47</sup>ISGlobal, Barcelona, Spain

<sup>48</sup>IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain

<sup>49</sup>Universitat Pompeu Fabra, Barcelona, Spain

<sup>50</sup>CIBER Epidemiología y Salud Pública, Madrid, Spain

<sup>51</sup>Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, MA, USA

<sup>52</sup>Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL, USA

<sup>53</sup>Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

<sup>54</sup>Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA

<sup>55</sup>International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland

<sup>56</sup>Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>57</sup>Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Copenhagen, Denmark

<sup>58</sup>Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany

<sup>59</sup>German Cancer Consortium, German Cancer Research Center Heidelberg, Germany

<sup>60</sup>Division of Preventive Oncology, German Cancer Research Center and National Center for Tumor Diseases, Heidelberg, Germany

<sup>61</sup>Humangenetik Tuebingen, Tuebingen, Germany

<sup>62</sup>University of Texas M.D. Anderson Cancer Center, Department of Genitourinary Medical Oncology, Houston, TX, USA

<sup>63</sup>Departments of Epidemiology & Population Health and of Medicine, Division of Oncology, Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA

<sup>64</sup>Department of Genetics, Portuguese Oncology Institute of Porto, Porto, Portugal

<sup>65</sup>Biomedical Sciences Institute, University of Porto, Porto, Portugal

<sup>66</sup>Cancer Genetics Group, IPO-Porto Research Center, Portuguese Oncology Institute of Porto, Porto, Portugal

<sup>67</sup>Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA, USA

<sup>68</sup>Ghent University, Faculty of Medicine and Health Sciences, Basic Medical Sciences, Ghent, Belgium

<sup>69</sup>Department of Surgery, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

<sup>70</sup>Department of Urology, University of Washington, Seattle, WA, USA

<sup>71</sup>Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>72</sup>Molecular Medicine Center, Department of Medical Chemistry and Biochemistry, Medical University of Sofia, Sofia, Bulgaria

<sup>73</sup>Department of Oncology, Cross Cancer Institute, University of Alberta, Edmonton, AB, Canada

<sup>74</sup>Division of Radiation Oncology, Cross Cancer Institute, Edmonton, AB, Canada

<sup>75</sup>Molecular Endocrinology Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium

<sup>76</sup>Division of Cancer Sciences, Manchester Cancer Research Centre, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, NIHR Manchester Biomedical Research Centre, Health Innovation Manchester, University of Manchester, Manchester, UK

<sup>77</sup>University of Surrey, Guildford, UK

<sup>78</sup>Genetic Oncology Unit, Complejo Hospitalario Universitario de Vigo, Instituto de Investigación Biomédica Galicia Sur, Vigo (Pontevedra), Spain

<sup>79</sup>Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands

<sup>80</sup>Exposome and Heredity, Faculté de Médecine, Université Paris-Saclay, Villejuif, France

<sup>81</sup>Clinical Gerontology Unit, University of Cambridge, Cambridge, UK

<sup>82</sup>Division of Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA

<sup>83</sup>George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT, USA

<sup>84</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

<sup>85</sup>Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>86</sup>Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

<sup>87</sup>Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

<sup>88</sup>International Epidemiology Institute, Rockville, MD, USA

<sup>89</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.

Funding and acknowledgements information for the PRACTICAL consortium, CRUK, BPC3, CAPS and PEGASUS are in the [Supplementary material](#), available as [Supplementary data](#) at *IJE* online.

## Ethics approval

Each individual study obtained ethical approval, therefore additional ethical approval for this study was not required.

## Data availability

Studies pooled by the EHNBPCCG are not owned by the writing group and so are not available from this consortium. Individual studies may be contacted to request access to their data. PRACTICAL consortium data are available upon request, see [<http://practical.icr.ac.uk/blog/>] for further details.

## Supplementary data

[Supplementary data](#) are available at *IJE* online.

## Author contributions

Author contributions are available as a [Supplementary file](#) at *IJE* online.

## Funding

This work was supported by Cancer Research UK (grant numbers C8221/A19170 and C8221/A29017) to fund the centralized pooling, checking and data analysis. E.L.W. was supported by the Nuffield Department of Population Health Early Career Research Fellowship. A.P.C. is supported by a Cancer Research UK Population Research Fellowship (C60192/A28516) and by the World Cancer Research Fund (WCRF UK), as part of the World Cancer Research Fund International grant programme (2019/1953). A.T.B.C. was supported in part by the Intramural Research Program of the National Institutes of Health and the National Cancer Institute. CLUE is funded by the American Institute for Cancer

Research, National Institutes of Health Grants (IU01AG18033 and IU01CA86308). H.I.M.S. was supported by research grants from the National Health and Medical Research Council of Australia. J.A.C.C. was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. M.E.C. is supported by the US National Institutes of Health (Grant U01 CA164973). R.M.M. was supported by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. R.M.M. was also supported by a Cancer Research UK (C18281/A29019) programme grant (the Integrative Cancer Epidemiology Programme). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

## Acknowledgements

We thank all participants, researchers and support staff who made the study possible. CHDS would like to acknowledge the support of Alice Whittemore and David Feldman. CLUE thank Kathy J Helzlsouer for her contributions to the cohort and thank the staff and participants of the CLUE study for their important contributions. Cancer data were provided by the Maryland Cancer Registry, Center for Cancer Prevention and Control, Maryland Department of Health, with funding from the State of Maryland and the Maryland Cigarette Restitution Fund. The collection and availability of cancer registry data are also supported by the Cooperative Agreement NU58DP006333, funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services. The authors would like to thank Paul Appleby for past contributions to establishing the EHNBPCCG dataset and related coding. We additionally thank investigators from BUPA, CHS, EPIC Norfolk and MMAS for contributing data. **DISCLAIMER:** 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

None declared.

## References

1. Ferlay J, Ervik M, Lam F *et al*. *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer, 2018. <https://gco.iarc.fr/today> (16 May 2022, date last accessed).
2. Allard JB, Duan C. IGF-binding proteins: why do they exist and why are there so many? *Front Endocrinol (Lausanne)* 2018;**9**:117.
3. Hakuno F, Takahashi S-I. 40 YEARS OF IGF1: IGF1 receptor signaling pathways. *J Mol Endocrinol* 2018;**61**:T69–T86.
4. LeRoith D, Holly JMP, Forbes B. The insulin-like growth factors: ligands, binding proteins and receptors. *Mol Metab* 2021; **52**:101245.



5. Weroha SJ, Haluska P. The insulin-like growth factor system in cancer. *Endocrinol Metab Clin North Am* 2012;**41**:335–50.
6. Stankovic S, Day FR, Zhao Y, *et al.*; 23andMe Research Team. Elucidating the genetic architecture underlying IGF1 levels and its impact on genomic instability and cancer risk. *Wellcome Open Res* 2021;**6**:20.
7. Knuppel A, Fensom GK, Watts EL *et al.* Circulating insulin-like growth factor-I concentrations and risk of 30 cancers: prospective analyses in UK biobank. *Cancer Res* 2020;**80**:4014–21.
8. 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**: 2288–300.
9. 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.
10. Schumacher FR, Al Olama AA, Berndt SI *et al.*; Genetic Associations and Mechanisms in Oncology (GAME-ON)/ Elucidating Loci Involved in Prostate Cancer Susceptibility (ELLIPSE) Consortium. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* 2018;**50**:928–36.
11. 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**:3192–204.
12. Conti DV, Darst BF, Moss LC *et al.* Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet* 2021;**53**:65–75.
13. Giambartolomei C, Vukcevic D, Schadt EE *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;**10**:e1004383.
14. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016;**45**:1866–86.
15. Key TJ, Appleby PN, Allen NE, Reeves GK. Pooling biomarker data from different studies of disease risk, with a focus on endogenous hormones. *Cancer Epidemiol Biomarkers Prev* 2010;**19**:960–65.
16. 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**:585–94.
17. NealeLab. *UK Biobank GWAS Results* 2019. <http://www.nealelab.is/uk-biobank> (16 May 2022, date last accessed).
18. Burgess S, Bowden J, Integrating summarized data from multiple genetic variants in Mendelian randomization: bias and coverage properties of inverse-variance weighted methods. arXiv arXiv:151204486. 27 Nov 2015, preprint: not peer reviewed.
19. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the  $I^2$  statistic. *Int J Epidemiol* 2016;**45**: 1961–74.
20. Palmer TM, Sterne JA, Harbord RM *et al.* Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol* 2011;**173**:1392–403.
21. Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;**40**:755–64.
22. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology* 2017;**28**: 30–42.
23. Kamat MA, Blackshaw JA, Young R *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;**35**:4851–53.
24. 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**:693–98.
25. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017;**32**:377–89.
26. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nat Commun* 2020;**11**: 376.
27. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet* 2017;**13**:e1007081.
28. Holmes MV, Richardson TG, Ference BA, Davies NM, Davey Smith G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nat Rev Cardiol* 2021;**18**:435–53.
29. Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLOS Genet* 2021;**17**: e1009440.
30. Deng Y, Pan W. A powerful and versatile colocalization test. *PLoS Comput Biol* 2020;**16**:e1007778.
31. Zheng J, Haberland V, Baird D *et al.* Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet* 2020;**52**:1122–31.
32. Liu B, Gloudemans MJ, Rao AS, Ingelsson E, Montgomery SB. Abundant associations with gene expression complicate GWAS follow-up. *Nat Genet* 2019;**51**:768–69.
33. McKay JD, Hung RJ, Han Y *et al.*; SpiroMeta Consortium. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet* 2017;**49**:1126–32.
34. Bonilla C, Lewis SJ, Rowlands M-A *et al.*; the PRACTICAL Consortium. Assessing the role of insulin-like growth factors and binding proteins in prostate cancer using Mendelian randomization: Genetic variants as instruments for circulating levels. *Int J Cancer* 2016;**139**:1520–33.
35. Holmes MV, Richardson TG, Ference BA, Davies NM, Davey Smith G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nat Rev Cardiol* 2021;**18**:435–53.
36. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 2015;**181**:251–60.
37. 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**:2274–88.
38. Chan JM, Stampfer MJ, Ma J *et al*. Insulin-like growth factor-I (IGF-I) and IGF binding protein-3 as predictors of advanced-stage prostate cancer. *J Natl Cancer Inst* 2002;**94**:1099–106.
  39. Muti P, Quattrin T, Grant BJB *et al*. Fasting glucose is a risk factor for breast cancer. *Cancer Epidemiol Biomarkers Prev* 2002;**11**:1361–8.
  40. 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**: 341–53.
  41. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;**40**:304–14.