The Epidemiology of Hormones and Prostate Cancer



Eleanor Watts

Wolfson College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2019

Acknowledgements

I want to thank my supervisors Tim, Ruth and Aurora for their tireless support and encouragement, which has made this DPhil an enjoyable and engaging experience.

I also want to thank Ni, Raph and the Endogenous Hormones, Nutritional Biomarkers Prostate Cancer Collaborative Group co-authors for their contributions to this research. I'm also thankful to the Cancer Epidemiology Unit staff who have been incredibly welcoming and have fuelled this research with their knowledge, caffeine and cake. Particularly, the EPIC group and my fellow Room 5 DPhils for their camaraderie and sense of fun.

Thanks also to the OUWLRC team, particularly Chris O' and Jill for all the "character building", and for teaching me the value of hard work. Finally, I am hugely grateful to my family and friends for their love and support.

The work from this thesis was funded by Cancer Research UK (grant numbers C8221/A19170, C8221/A20986).

Abstract

The Epidemiology of Hormones and Prostate Cancer

Eleanor Watts, Wolfson College, University of Oxford, DPhil, Trinity term 2019.

Prostate cancer is the most common cancer in men in the UK. However, there are no well-established modifiable risk factors. Testosterone and insulin-like growth factor-I (IGF-I) are hormones which may be involved in prostate cancer risk, but the totality of the evidence is not clear. This thesis aimed to investigate the epidemiology and associations of circulating testosterone, IGF-I and other factors with prostate cancer risk.

The cross-sectional associations of a range of anthropometric, sociodemographic, and behavioural factors with circulating sex hormone and IGF concentrations were investigated using the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG), an international consortium of prospective studies. These analyses used individual participant data from up to 25 studies and 16,000 men free from prostate cancer and observed that age, adiposity and racial/ethnic group generally had large magnitudes of associations with these hormone concentrations. This enhances the understanding of how these factors may affect disease risk and may offer insights for prostate cancer prevention.

The association between free testosterone (the testosterone which is considered biologically active) and prostate cancer risk was investigated in the context of a possible threshold association. The EHNBPCCG contains individual participant data from 20 prospective studies including 7,000 prostate cancer cases and 12,000 controls with free testosterone data. Men with very low circulating free testosterone concentrations had a lower risk of prostate cancer, but possibly a higher risk of high-grade tumours. This analysis provides the first large-scale prospective epidemiological evidence of an association between circulating free testosterone concentrations and prostate cancer risk.

The possible associations of clinically low testosterone (diagnosis of Klinefelter's syndrome, testicular hypofunction and hypopituitarism) and clinically high IGF-I (acromegaly) with subsequent risk of prostate cancer diagnosis and mortality were examined using English national hospital admission and death records. Men diagnosed with any hypogonadal disease had a lower risk of prostate cancer mortality.

Men diagnosed with acromegaly had an increased risk of prostate cancer diagnosis and possibly prostate cancer mortality. These results support the roles of both testosterone and IGF-I in prostate cancer risk.

Possible haematological risk factors for prostate cancer were investigated in the UK Biobank in 210,000 men (5,700 prostate cancer cases). Higher red blood cell and platelet counts, and lower measures of red blood cell size were associated with an increased prostate cancer risk. Higher white blood cell and neutrophil counts were associated with an increased risk of prostate cancer mortality. These findings were consistent with the possible roles of testosterone, inflammation and/or infection in prostate cancer risk.

In conclusion, research from this thesis supports the roles of testosterone, IGF-I and infection/inflammation in prostate cancer risk and identified potential anthropometric, sociodemographic and behavioural factors which may affect prostate cancer risk through alterations in these physiological characteristics.

Contents

List of Figures **xi**

List of Tables **xv**

List of Abbreviations **xvii**

1 Introduction **1**
- 1.1 Rationale . . . 1
- 1.2 Aims . . . 2
- 1.3 Outline of the thesis . . . 2
- 1.4 The role of the author in the preparation of this thesis . . . 4
- 1.5 Publications arising from this thesis . . . 5

2 Prostate cancer epidemiology and hormones **6**
- 2.1 Introduction . . . 7
- 2.2 Trends in prostate cancer incidence and mortality . . . 7
  - 2.2.1 Geographic trends . . . 7
  - 2.2.2 Temporal trends . . . 9
- 2.3 Prostate-specific antigen testing . . . 11
- 2.4 Classification of prostate cancer . . . 12
- 2.5 Established risk factors . . . 17
  - 2.5.1 Age . . . 17
  - 2.5.2 Family history of prostate cancer . . . 18
  - 2.5.3 Genetics . . . 18
  - 2.5.4 Racial/ethnic group . . . 19
- 2.6 Putative risk factors . . . 21
  - 2.6.1 Height . . . 21
  - 2.6.2 Body fat . . . 22
  - 2.6.3 Co-morbidities . . . 24
  - 2.6.4 Inflammation and infection . . . 25
  - 2.6.5 Diet . . . 25
- 2.7 Sex hormones and insulin-like growth factors . . . 29
  - 2.7.1 Sex hormones . . . 29

2.7.2 Insulin-like growth factors . . . 35
2.7.3 Hormone action in the prostate . . . 38
2.7.4 Hormones and prostate cancer . . . 46
2.8 Conclusion . . . 49

3 Studies used in this thesis 50
3.1 Introduction . . . 51
3.2 The Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group . . . 51
3.2.1 Introduction . . . 51
3.2.2 Key contributors to the EHNBPCCG dataset and analysis . 52
3.2.3 Study identification . . . 52
3.2.4 Participating studies . . . 52
3.2.5 Ethical approval . . . 60
3.3 English national health records . . . 61
3.3.1 Introduction . . . 61
3.3.2 Hospital Episode Statistics Admitted Patient Care (HES APC) 61
3.3.3 Ethical approval . . . 63
3.4 UK Biobank . . . 65
3.4.1 Introduction . . . 65
3.4.2 Design and methods . . . 65
3.4.3 Ethical approval . . . 69
3.4.4 Follow-up . . . 70

4 Cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating sex hormone concentrations in men 71
4.1 Introduction . . . 72
4.1.1 Aims . . . 73
4.2 Subjects and methods . . . 73
4.2.1 Data collection . . . 73
4.2.2 Data processing . . . 80
4.2.3 Statistical analysis . . . 82
4.2.4 Multiple testing . . . 84
4.2.5 Statistical software . . . 84
4.3 Results . . . 84
4.3.1 Correlations between hormones . . . 84
4.3.2 Age at blood collection . . . 86
4.3.3 Adiposity . . . 89
4.3.4 Height . . . 93

4.3.5 Smoking . . . 93
4.3.6 Alcohol . . . 93
4.3.7 Ethnic/racial group . . . 97
4.3.8 Time of blood collection . . . 97
4.3.9 Sociodemographic factors and health . . . 97
4.3.10 Heterogeneity by study and assay . . . 103
4.3.11 Sensitivity analyses . . . 103
4.4 Discussion . . . 104
4.4.1 Correlations between hormones . . . 104
4.4.2 Age . . . 104
4.4.3 Adiposity . . . 105
4.4.4 Height . . . 106
4.4.5 Smoking . . . 106
4.4.6 Alcohol . . . 106
4.4.7 Racial/ethnic group . . . 107
4.4.8 Heterogeneity between studies . . . 107
4.4.9 Calculation of free testosterone and estradiol . . . 107
4.4.10 Strengths and limitations . . . 108
4.5 Conclusion . . . 109

5 Cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating IGF and IGFBP concentrations in men 110
5.1 Introduction . . . 111
5.1.1 Aims . . . 111
5.2 Subjects and methods . . . 112
5.2.1 Data collection . . . 112
5.2.2 Data processing . . . 118
5.2.3 Statistical analysis . . . 118
5.2.4 Multiple testing . . . 120
5.3 Results . . . 120
5.3.1 Correlations between IGF concentrations . . . 120
5.3.2 Age at blood collection . . . 122
5.3.3 Adiposity . . . 124
5.3.4 Height . . . 124
5.3.5 Smoking . . . 129
5.3.6 Alcohol . . . 129
5.3.7 Ethnic/racial group . . . 129
5.3.8 Sociodemographic factors and health . . . 133

5.3.9 Tests for heterogeneity and sensitivity analysis 133
5.3.10 Further analyses of IGFs and sex hormones 139
5.4 Discussion 141
5.4.1 Correlations between hormones 141
5.4.2 Age 141
5.4.3 Adiposity 142
5.4.4 Height 142
5.4.5 Smoking 143
5.4.6 Alcohol 143
5.4.7 Racial/ethnic group 143
5.4.8 Sociodemographic factors and health 144
5.4.9 Heterogeneity by study 144
5.4.10 Correlations between IGFs and sex hormones 144
5.4.11 Strengths and limitations 145
5.5 Conclusion 146

6 Analysis of circulating free testosterone and prostate cancer risk 147
6.1 Introduction 148
6.1.1 Context 148
6.1.2 Biological threshold 149
6.1.3 Aims 150
6.2 Subjects and methods 151
6.2.1 Data collection 151
6.2.2 Inclusion criteria 151
6.2.3 Study design and matching 151
6.2.4 Data processing 157
6.2.5 Statistical analysis 157
6.2.6 Tests for heterogeneity 158
6.2.7 Sensitivity analysis 159
6.2.8 Statistical software 159
6.3 Results 160
6.3.1 Associations between circulating sex hormones and prostate cancer risk 165
6.3.2 Heterogeneity by study 167
6.3.3 Heterogeneity by tumour and participant characteristics . . 167
6.3.4 Other assay measures 170
6.3.5 Sensitivity analysis 170
6.4 Discussion 172
6.4.1 Main findings 172

6.4.2 Biological plausibility . . . 172
6.4.3 Heterogeneity by tumour grade . . . 172
6.4.4 Detection bias . . . 174
6.4.5 Co-morbidities . . . 174
6.4.6 IGF-I at blood collection . . . 175
6.4.7 Free hormone hypothesis . . . 175
6.4.8 Strengths and limitations . . . 176
6.5 Conclusion . . . 178

7 Hormone-related diseases and prostate cancer risk 179
7.1 Introduction . . . 180
7.1.1 Hormone-related disorders . . . 180
7.1.2 Aims . . . 183
7.2 Methods . . . 184
7.2.1 Study design and population . . . 184
7.2.2 Statistical analysis . . . 189
7.2.3 Statistical software . . . 190
7.3 Results . . . 190
7.3.1 Reference cohort . . . 190
7.3.2 Acromegaly cohort . . . 190
7.3.3 Hypogonadal cohort . . . 193
7.3.4 Further analyses . . . 193
7.4 Discussion . . . 198
7.4.1 Acromegaly cohort . . . 198
7.4.2 Hypogonadal cohort . . . 198
7.4.3 Detection bias . . . 200
7.4.4 Strengths and limitations . . . 201
7.5 Conclusion . . . 203

8 Haematological parameters and prostate cancer risk in UK Biobank 204
8.1 Introduction . . . 205
8.1.1 Determinants of red blood cells . . . 205
8.1.2 Determinants of platelets and white blood cells . . . 205
8.1.3 Modifiable risk factors for prostate cancer . . . 205
8.1.4 Aim . . . 206
8.2 Methods . . . 206
8.2.1 UK Biobank . . . 206
8.2.2 Exclusion criteria . . . 210
8.2.3 Statistical analysis . . . 211
8.2.4 Further analyses . . . 213

8.2.5 Statistical software 214
8.3 Results 215
8.3.1 Correlation between blood indices 219
8.3.2 Associations between haematological indices and prostate cancer incidence and mortality 221
8.3.3 Further analyses 229
8.4 Discussion 236
8.4.1 Associations with red blood cell indices 236
8.4.2 Associations with platelet and white blood cell indices . . . 236
8.4.3 Diet and other factors 237
8.4.4 Reverse causation 237
8.4.5 Strengths and limitations 238
8.4.6 Conclusion 239

9 Summary and recommendations for future research 240
9.1 Introduction 241
9.2 Summary of findings 241
9.2.1 Findings across analyses and contribution to the literature . 243
9.3 Methodological considerations 246
9.3.1 The role of chance 246
9.3.2 The role of bias 247
9.3.3 The role of confounding 254
9.3.4 Summary of limitations 254
9.4 Priorities for future research 256
9.4.1 Hormones in healthy and malignant prostate tissue 256
9.4.2 Triangulation 256
9.4.3 Mediation analysis 259
9.4.4 Aggressive tumour subtypes 259
9.4.5 Collaboration and data pooling 260
9.4.6 Implications for prostate cancer prevention 260
9.5 Conclusion 261

Appendices

A Appendix 1 263
A.1 Alternative free hormone calculations 263
A.1.1 Free testosterone 263
A.1.2 Free estradiol 263

References 265

List of Figures

2.1 Estimated age-standardised global incidence rates of prostate cancer in males, 2018 8
2.2 Estimated age-standardised global mortality rates of prostate cancer in males, 2018 9
2.3 Age-standardised prostate cancer incidence and mortality rate per 100,000 US men by year 10
2.4 Prostate cancer stage distribution at diagnosis in US men, 2005-2014 11
2.5 Prostate cancer incidence rates per 100,000 per year in US men by age at diagnosis, 2005-2014 18
2.6 Simplified schematic of the biosynthesis pathway of sex hormones . 31
2.7 Simplified schematic for the main secretory agents and regulatory mechanisms of sex steroid hormones 33
2.8 Somatotropic axis 36
2.9 Prostate tissue structure 39
2.10 Schematic diagram of DHT synthesis in the prostate 42
2.11 The androgen saturation model of intra-prostatic androgen signalling 43
2.12 IGF signalling in the prostate 45

3.1 Medical Certificate of Cause of Death 64

4.1 Participant selection chart 77
4.2 Partial correlation coefficients between concentrations of male sex hormones and SHBG 85
4.3 Relative geometric mean concentrations of male sex hormones by age 87
4.4 Relative geometric mean concentrations of testosterone by age, stratified by time of blood collection 88
4.5 Relative geometric mean concentrations of male sex hormones by BMI 90
4.6 Relative geometric mean concentrations of male sex hormones by waist circumference 91
4.7 Relative geometric mean concentrations of male sex hormones by WHR 92
4.8 Relative geometric mean concentrations of male sex hormones by height. 94

4.9 Relative geometric mean concentrations of male sex hormones by smoking status . . . 95
4.10 Relative geometric mean concentrations of male sex hormones by alcohol consumption . . . 96
4.11 Relative geometric mean concentrations of male sex hormones by racial/ethnic group . . . 98
4.12 Relative geometric mean concentrations of male sex hormones by time of blood collection . . . 99
4.13 Relative geometric mean concentrations of male sex hormones by education status . . . 100
4.14 Relative geometric mean concentrations of male sex hormones by marriage status . . . 101
4.15 Relative geometric mean concentrations of male sex hormones by family history of prostate cancer . . . 102

5.1 Participant selection chart. . . . 116
5.2 Partial correlation coefficients IGFs and IGFBPs . . . 122
5.3 Relative geometric mean concentrations of male IGFs and IGFBPs by age . . . 123
5.4 Relative geometric mean concentrations of male IGFs and IGFBPs by BMI . . . 125
5.5 Relative geometric mean concentrations of male IGFs and IGFBPs by waist circumference . . . 126
5.6 Relative geometric mean concentrations of male IGFs and IGFBPs by WHR . . . 127
5.7 Relative geometric mean concentrations of male IGFs and IGFBPs by height . . . 128
5.8 Relative geometric mean concentrations of male IGFs and IGFBPs by smoking status . . . 130
5.9 Relative geometric mean concentrations of male IGFs and IGFBPs by alcohol consumption . . . 131
5.10 Relative geometric mean concentrations of male IGFs and IGFBPs by racial/ethnic group . . . 132
5.11 Relative geometric mean concentrations of IGFs and IGFBPs in males by marriage status . . . 134
5.12 Relative geometric mean concentrations of IGFs and IGFBPs in males by education status . . . 135
5.13 Relative geometric mean concentrations of IGFs and IGFBPs in males by family history of prostate cancer . . . 136
5.14 Relative geometric mean concentrations of IGFBP-1 by fasting status 137

5.15 Relative geometric mean concentrations of IGFBP-2 by fasting status 138
5.16 Partial correlation coefficients between male sex hormones, SHBG, IGFs and IGFBPs 140

6.1 Geometric mean (95% CIs) total testosterone and SHBG concentrations by study-specific tenths of free testosterone 164
6.2 Odds ratio (95% CIs) of prostate cancer risk by study-specific tenths of hormone concentrations 166
6.3 Odds ratio (95% CIs) of prostate cancer risk by study-specific tenths free testosterone concentrations 167
6.4 Odds ratio (95% CIs) of prostate cancer risk by study-specific free testosterone concentrations (1^{st} vs. 2^{nd}–10^{th}) by study 168
6.5 Odds ratio (95% CIs) of prostate cancer risk by study-specific tenths of free testosterone concentrations (1^{st} vs. 2^{nd}–10^{th}), by participant characteristics 169
6.6 Geometric mean PSA and IGF-I concentrations by study-specific tenth of free testosterone 171

7.1 Schematic of cohort selection criteria 188
7.2 Hazard ratio of prostate cancer diagnosis and mortality 192
7.3 Hazard ratio of prostate cancer diagnosis in men diagnosed with acromegaly, by age at cohort entry and time interval 195
7.4 Hazard ratio of prostate cancer diagnosis in men diagnosed with hypogonadal-associated diseases, by age at cohort entry and time interval 196
7.5 Hazard ratio of prostate cancer diagnosis, by length of follow-up . . 197

8.1 Participant selection chart 211
8.2 Partial correlation matrix of haematological indices 220
8.3 Hazard ratio (95% CIs) of prostate cancer diagnosis by fifths of red blood cell indices 222
8.4 Hazard ratio (95% CIs) of prostate cancer mortality by fifths of red blood cell indices 223
8.5 Hazard ratio (95% CIs) of prostate cancer diagnosis by fifths of platelet indices 224
8.6 Hazard ratio (95% CIs) of prostate cancer mortality by fifths of platelet indices 225
8.7 Hazard ratio (95% CIs) of prostate cancer diagnosis by fifths of white blood cell indices 226
8.8 Hazard ratio (95% CIs) of prostate cancer mortality by fifths of white blood cell indices 227

8.9 Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by age at recruitment 232
8.10 Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by category of BMI 233
8.11 Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by smoking status 234
8.12 Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by diabetes status 235

9.1 The Mendelian randomization approach to investigating free testosterone and prostate cancer risk 258

List of Tables

2.1 TNM tumour staging system for prostate cancer 14
2.2 AJCC Prostate cancer staging. 15

3.1 Collaborating studies in the Endogenous Hormones, Nutritional Biomarkers Prostate Cancer Collaborative Group dataset. 56

4.1 Assay methods and geometric mean sex hormone and SHBG concentrations. 74
4.2 Assay methods and geometric mean sex hormone and SHBG concentrations (continued). 75
4.3 Assay methods and geometric mean sex hormone and SHBG concentrations (continued). 76
4.4 Participant characteristics by study. 79

5.1 Assay methods and geometric mean IGF concentrations. 113
5.2 Assay methods and geometric mean IGFBP concentrations. 114
5.3 Participant characteristics by studies with IGF data 117

6.1 Sample population, recruitment and case ascertainment method by study. 153
6.2 Criteria used by individual studies to match case patients and control subjects. 155
6.3 Assay methods and geometric mean testosterone and SHBG concentrations by study. 156
6.4 Characteristics of patients diagnosed with prostate cancer by study 161
6.5 Geometric mean free testosterone concentrations by study-specific tenths (1^{st} vs 2^{nd}-10^{th}) and 1^{st} decile cut-point 162
6.6 Characteristics of prostate cancer cases and controls by study-specific tenths of calculated free testosterone concentration. 163
6.7 Low free testosterone by study-specific tenths of total testosterone and SHBG. 165

7.1 Identification of prostate cancer by HES APC and death records. . 187
7.2 Participant characteristics at cohort entry 191

8.1 Summary description of haematological parameters and measurement/calculation methods by UK Biobank. 209
8.2 Baseline characteristics and blood indices of men who remained free from prostate cancer and men who developed prostate cancer 216
8.3 Baseline characteristics by red blood cell, platelet and white blood cell counts at initial assessment 217
8.4 Median blood indices at initial assessment and first resurvey and regression dilution ratios. 218
8.5 Hazard ratios for prostate cancer diagnosis and mortality per 1 SD increase (95% CI) in blood parameters with and without correction for regression dilution bias 228
8.6 Hazard ratios for prostate cancer diagnosis per 1 SD increase in blood parameters by time to diagnosis. 230
8.7 Hazard ratios for prostate cancer diagnosis per 1 SD increase in blood parameters by age at diagnosis. 231

List of Abbreviations

A & E . . . . . Accident and Emergency

ACTH . . . . . Adrenocorticotrophic hormone

A-diol-G . . . Androstanediol glucuronide

AJCC . . . . . American Joint Committee on Cancer

ALS . . . . . Acid-labile subunit

ATBC . . . . . Alpha Tocopherol, Beta Carotene Prevention Trial

BLSA . . . . . Baltimore Longitudinal Study of Aging

BMI . . . . . Body mass index

BPH . . . . . Benign prostatic hyperplasia

BUPA . . . . . British United Provident Association

CAP . . . . . Cluster randomised triAl of PSA testing for Prostate cancer

CARET . . . . Carotene and Retinol Efficacy Trial

CHDS . . . . . Child Health and Development Studies

CHS . . . . . Cardiovascular Health Study

CI . . . . . Confidence interval

CLUE . . . . . Give us a Clue to Cancer

CRH . . . . . Corticotrophin releasing hormone

DHEA . . . . . Dehydroepiandrosterone

DHEAS . . . . Dehydroepiandrosterone sulphate

DHT . . . . . Dihydrotestosterone

EDTA . . . . . Ethylenediaminetetraacetic acid

EHNBPCCG Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group

ELISA . . . . . Enzyme-linked immunosorbent assay

EPIC . . . . . European Prospective Investigation into Cancer and Nutrition

ERSPC . . . . European Randomised Study of Prostate Cancer

FMC . . . . . Finnish Mobile Clinic

FSH . . . . . Follicle stimulating hormone

GAME-ON . . Genetic Associations and Mechanisms in Oncology

GH . . . . . Growth hormone

GnRH . . . . . Gonadotropin releasing hormone

GP . . . . . General Practice

GWAS . . . . . Genome-wide association study

HES APC . . Hospital Episode Statistics Admitted Patient Care

HGPIN . . . . High-grade prostatic intraepithelial neoplasia

HHS . . . . . Helsinki Heart Study

HIMS . . . . . Health In Men Study

HPA . . . . . Hypothalamic-pituitary-adrenal

HPFS . . . . . Health Professionals Follow-up Study

HPG . . . . . Hypothalamus-pituitary-gonadal

HR . . . . . Hazard ratio

HSD . . . . . Hydroxysteroid dehydrogenase

HST . . . . . Hydroxysteroid sulfotransferase

ICD-10 . . . . International Classification of Disease 10^{th} revision

ICD-9 . . . . . International Classification of Disease 9^{th} revision

IGF . . . . . Insulin-like growth factor

IGF-1R . . . . Type 1 insulin receptor

IGF-2R . . . . Type 2 insulin receptor

IGFBP . . . . Insulin-like growth factor binding protein

IMD . . . . . Index of multiple deprivation

IQR . . . . . Interquartile range

IU . . . . . International unit

JCCS . . . . . Japan Collaborative Cohort Study

JHCS . . . . . Japan-Hawaii Cancer Study

JPHC . . . . . Japan Public Health Center-based prospective study

KPMCP . . . Kaiser Permanente Medical Care Program

LH . . . . . Luteinising hormone

MAPK/ERK Mitogen-activated protein kinase/extracellular signal-regulated kinases

MCCS . . . . . Melbourne Collaborative Cohort Study

MCH . . . . . Mean corpuscular haemoglobin

MCHC . . . . Mean corpuscular haemoglobin concentration

MCV . . . . . Mean corpuscular volume

MEC . . . . . Multiethnic Cohort

MMAS . . . . Massachusetts Male Aging Study

MSCV . . . . . Mean sphered cell volume

NHS . . . . . National Health Service

NSHDC . . . . Northern Sweden Health and Disease Cohort

ONS . . . . . Office of National Statistics

OR . . . . . Odds ratio

PCPT . . . . . Prostate Cancer Prevention Trial

PHS . . . . . Physicians' Health Study

PI3K . . . . . Phosphoinositide 3-kinase

PLCO . . . . . Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

PRACTICAL Prostate Cancer Association Group to Investigate Cancer Associated Alterations to the Genome

ProtecT . . . . Prostate Testing for Cancer and Treatment

PSA . . . . . Prostate-specific antigen

RBS . . . . . Rancho Bernado Study

REDUCE . . . Reduction by Dutasteride of Prostate Cancer Events Trial

RR . . . . . Relative risk

SD . . . . . Standard deviation

SELECT . . . Selenium and Vitamin E Cancer Prevention Trial

SHBG . . . . . Sex hormone binding globulin

SNP . . . . . Single nucleotide polymorphism

SU.VI.MAX . SUpplémentation en VItamines et Minéraux AntioXydants

TNM . . . . . Tumour-Node-Metastasis

TRT . . . . . Testosterone replacement therapy

UK BiLEVE . UK Biobank Lung Exome Variant Evaluation

USPSTF . . . US Preventative Services Task Force

WCRF . . . . World Cancer Research Fund

WHR . . . . . Waist-to-hip ratio

1 Introduction

Contents

1.1 Rationale 1
1.2 Aims 2
1.3 Outline of the thesis 2
1.4 The role of the author in the preparation of this thesis 4
1.5 Publications arising from this thesis 5

1.1 Rationale

Prostate cancer is the second most commonly diagnosed cancer in men worldwide and the fifth leading cause of cancer death [1, 2]. Although there is substantial international variation in prostate cancer incidence, there are no well-established preventable risk factors [3].

Testosterone and insulin-like growth factor-I (IGF-I) are hormones which may be involved in prostate cancer development and progression. Animal models [4, 5], genetic evidence [6, 7] and two randomised controlled trials [8–10] support the role of androgens in prostate cancer development and testosterone deprivation therapy is the mainstay treatment for advanced prostate cancer [11, 12]. However, there is no prospective epidemiological evidence of an association between circulating

testosterone concentrations and subsequent prostate cancer risk [13].

Associations between circulating IGF-I concentrations and prostate cancer risk are generally consistent across prospective epidemiological studies [14], animal models [15] and clinical studies [16, 17] and are therefore considered more robust. However, the associations between IGF-I and prostate cancer mortality in prospective epidemiological studies are less clear [14, 18].

1.2 Aims

Analyses in this thesis aimed to investigate the epidemiology and associations of circulating testosterone, IGF-I and other blood characteristics with prostate cancer risk.

1.3 Outline of the thesis

- **Chapter 1:** introduces the thesis and summarises the thesis rationale and the aims that will be explored in Chapters 4 to 8.

- **Chapter 2:** describes prostate cancer epidemiology, including global trends in incidence and mortality, and discusses the established and putative risk factors for prostate cancer. This chapter also provides an overview of sex hormones and IGFs including steroidogenesis, regulation, biological activity in the prostate and hypothesised associations with prostate cancer risk.

- **Chapter 3:** describes the datasets analysed in this thesis. These include the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG), English national Hospital Episode Statistics Admitted Patient Care (HES APC) and national mortality records, and the UK Biobank.

- **Chapter 4:** investigates the cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating sex hormones in men using the EHNBPCCG dataset.

- **Chapter 5:** investigates the cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating IGF and IGFBP concentrations in men using the EHNBPCCG dataset.

- **Chapter 6:** investigates the association of low free testosterone concentration with incident prostate cancer using the EHNBPCCG dataset.

- **Chapter 7:** investigates the associations of diseases which are characterised by clinically high IGF-I or clinically low testosterone concentrations with risk of prostate cancer diagnosis and mortality using HES APC and national mortality records.

- **Chapter 8:** investigates the associations of complete blood cell count and reticulocyte indicies with prostate cancer diagnosis and mortality using the UK Biobank.

- **Chapter 9:** summarises and interprets the main findings of this thesis, discusses the strengths and limitations of this research, and outlines priorities for future research.

1.4 The role of the author in the preparation of this thesis

The work from this thesis was conducted under the supervision of Prof. Timothy Key, Assoc. Prof. Ruth Travis, and Dr. Aurora Perez-Cornago. Timothy Key and Ruth Travis secured the funding for the work from Cancer Research UK. After the conception of the main aim for the DPhil, I have been involved in all aspects of the research and have led the data analysis, interpretation, and write-up.

The EHNBPCCG pooled dataset contains individual participant data (which is generally published), as such this group had no role in the individual study designs or assays, with the exception of the case-control studies nested in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort.

I wrote all the chapters contained in this thesis. However, co-authors have read and provided feedback on the manuscripts which were submitted for publication in Chapters 4-7. Dr. Jaimal Kothari, a clinical haematologist, was consulted to aid the interpretation of the findings in Chapter 8.

I conducted all statistical analyses for Chapters 4-6 and 8. For Chapter 7, I developed the analysis plan and Raphael Goldacre conducted the statistical analysis because the datasets are held in the Unit of Health Care Epidemiology and making these available in the Cancer Epidemiology Unit would have expended substantial resources. Raphael Goldacre also commented on the manuscript which was submitted for publication. All tables and figures in this thesis were created by me.

1.5 Publications arising from this thesis

At the time of writing, five manuscripts have been published, for which I am the lead author:

- **Chapter 4:** The cross-sectional associations of sex hormones has been published in PLoS ONE (Watts *et al.*, 2017 PLoS One. Dec 27;12(12):e0187741). Available from: `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0187741`

- **Chapter 5:** The cross-sectional associations of IGFs and IGFBPs has been published in the International Journal of Cancer (Watts *et al.*, 2019 Int J Cancer 145(12):3244-3256. doi: 10.1002/ijc.32276). Available from: `https://onlinelibrary.wiley.com/journal/10970215`

- **Chapter 6:** Low free testosterone and prostate cancer risk has been published in European Urology (Watts *et al.*, 2018 Euro Urol Nov;74(5):585-594). Available from: `https://www.europeanurology.com/article/S0302-2838(18)30546-3/fulltext`

- **Chapter 7:** Hormone-related diseases and prostate cancer risk has been published in the International Journal of Cancer (Watts *et al.*, 2019 Int J Cancer 147(3):803-810. doi: 10.1002/ijc.32808). Available from: `https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.32808`

- **Chapter 8:** Haematological parameters and prostate cancer risk in UK Biobank has been published in Cancer Epidemiology and Prevention Biomarkers (Watts *et al.*, 2020 Cancer Epidemiol Biomarkers Prev 29(8):1615-1626. doi: 10.1158/1055-9965.EPI-19-1525). Available from: `https://cebp.aacrjournals.org/content/29/8/1615.long`

2

Prostate cancer epidemiology and hormones

Contents

2.1 Introduction 7
2.2 Trends in prostate cancer incidence and mortality . . 7
2.2.1 Geographic trends 7
2.2.2 Temporal trends 9
2.3 Prostate-specific antigen testing 11
2.4 Classification of prostate cancer 12
2.5 Established risk factors 17
2.5.1 Age 17
2.5.2 Family history of prostate cancer 18
2.5.3 Genetics 18
2.5.4 Racial/ethnic group 19
2.6 Putative risk factors 21
2.6.1 Height 21
2.6.2 Body fat 22
2.6.3 Co-morbidities 24
2.6.4 Inflammation and infection 25
2.6.5 Diet 25
2.7 Sex hormones and insulin-like growth factors 29
2.7.1 Sex hormones 29
2.7.2 Insulin-like growth factors 35
2.7.3 Hormone action in the prostate 38
2.7.4 Hormones and prostate cancer 46
2.8 Conclusion 49

2.1 Introduction

Prostate cancer is the second most commonly diagnosed cancer in men worldwide and the fifth leading cause of cancer death [1, 2]. Although there is substantial international variation in prostate cancer incidence, there are no well-established preventable risk factors. Testosterone and insulin-like growth factor-I (IGF-I) are hormones which are modifiable and may be associated with prostate cancer aetiology and development [14, 19].

This chapter summarises the global trends in prostate cancer diagnosis and mortality, outlines established and putative risk factors, and introduces sex hormone and IGF biosynthesis and regulation, as well as their prostate-specific biological action and their hypothesised relationships with prostate cancer risk.

2.2 Trends in prostate cancer incidence and mortality

2.2.1 Geographic trends

Globocan estimates that 1.3 million new cases of prostate cancer were diagnosed in 2018 and 360,000 men died of prostate cancer [1, 2]. Prostate cancer is most common in Western countries, with an age-standardised incidence rate of 86.4 and 85.7 per 100,000 male population in Australia/New Zealand and Northern Europe, respectively. Incidence rates are lowest in Asian populations, with estimated age-standardised incidence of 12.7 and 5.0 per 100,000 men in South-Eastern and South-Central Asia, respectively [1, 2] (Figure 2.1).

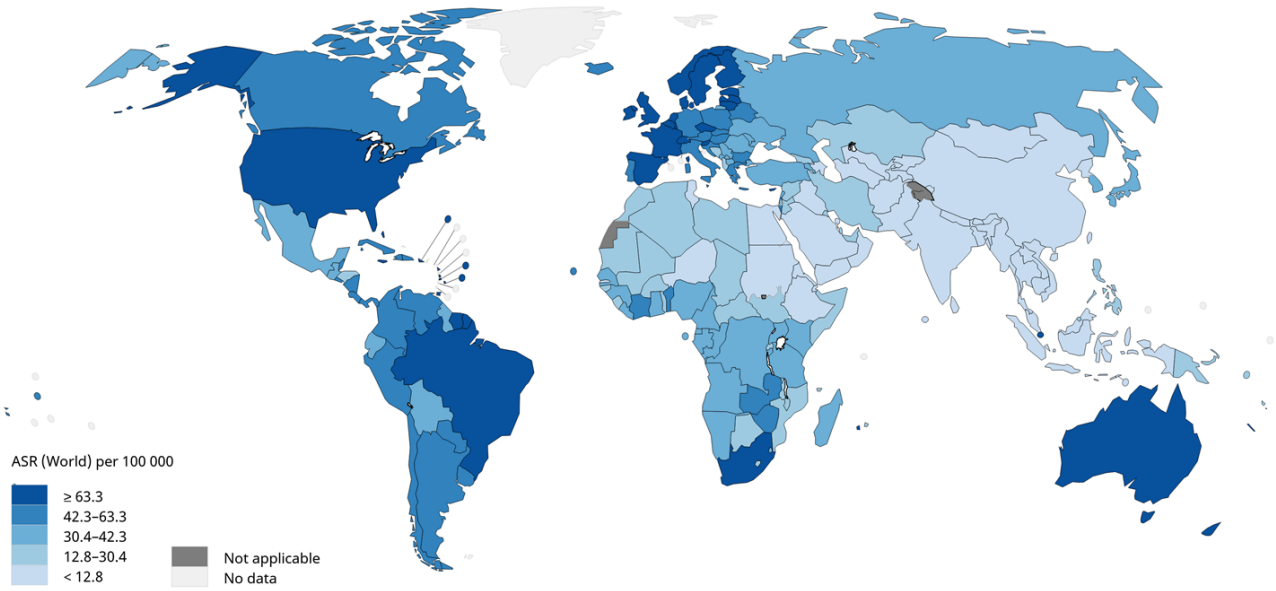


Figure 2.1: Estimated age-standardised global incidence rates of prostate cancer in males, 2018.
Data source: [1, 2].

There is less variation in prostate cancer mortality (Figure 2.2). Mortality is highest amongst predominantly black populations; age-standardised mortality is estimated to be 26.8 and 25.4 per 100,000 male population in Southern Africa and the Caribbean, respectively [1, 2]. Rates are lowest in Asian populations: 3.3 and 4.7, per 100,000 men, in South-Central and Eastern Asia, respectively [1, 2]. Despite reporting the highest incident prostate cancer cases, age-standardised mortality across Europe, North America and Oceania is reported to be between 7.7 to 13.5 per 100,000 [1, 2].

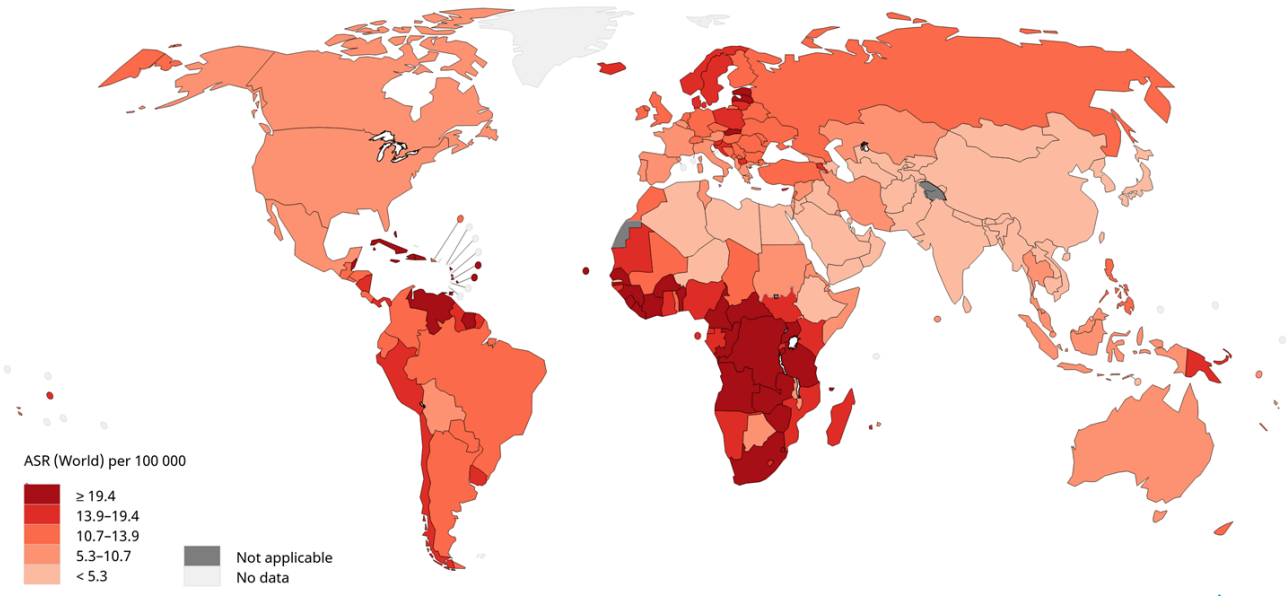


Figure 2.2: Estimated age-standardised global mortality rates of prostate cancer in males, 2018.
Data source: [1, 2].

2.2.2 Temporal trends

The greatest contributor to the geographic variation in prostate cancer incidence may be differences in national screening policies, particularly the prevalence of prostate-specific antigen (PSA) testing [20] (see below, Section 2.3). PSA testing became clinically available between the late-1980s and early 1990s and is used both as a diagnostic and a prostate cancer screening tool [21, 22]. The prominence of PSA testing has led to large increases in prostate cancer diagnosis in countries where the practice of testing has been common, such as the in the USA [23] (Figure 2.3), where the majority of prostate cancer is detected at the localised stage (Figure 2.4). However, PSA testing rates in the US are now declining [24], which may partly explain the more recent decreases in age-standardised prostate cancer diagnosis (Figure 2.3).

In other high income countries, such as the UK, PSA testing has had a slower uptake which has led to a more steady increase in prostate cancer diagnosis [25]. Whilst in lower income countries, diagnosis is generally increasing more gradually with lower usage of PSA testing [26].

Prostate cancer mortality is generally decreasing in high income countries [26]. This is attributable to improved treatment methods, access to healthcare, and possibly early detection. Mortality rates are increasing in many countries across Central and Eastern Europe, Asia and Africa [26]. Global variations in prostate cancer mortality rates are likely driven by differences in access to healthcare, treatment practices, and death certificate reporting [27, 28]. Clinical studies also suggest that prostate tumours may be more aggressive in men of African ancestry [29] (see Section 2.5.4).

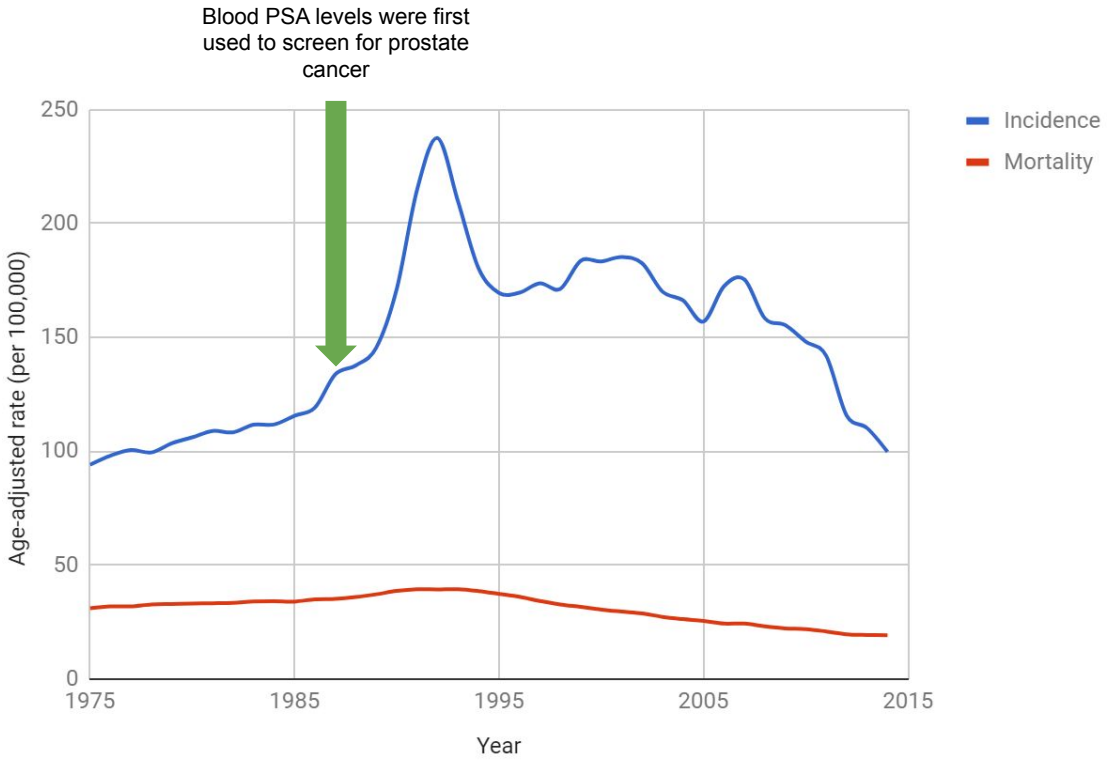


Figure 2.3: Age-standardised prostate cancer incidence and mortality rate per 100,000 US men by year.
Data source: Incidence and mortality rates [30]; PSA date [31].
Abbreviations: PSA=prostate-specific antigen.

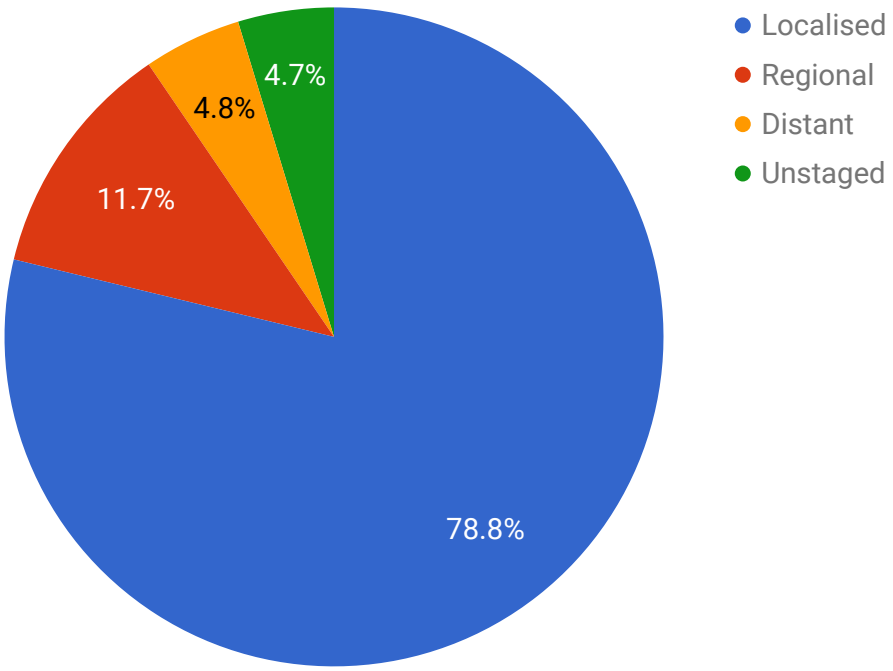


Figure 2.4: Prostate cancer stage distribution at diagnosis in US men, 2005-2014. Data source: [30].

2.3 Prostate-specific antigen testing

In contrast with breast, cervix and colorectal cancer, there is no formal screening programme for prostate cancer. PSA is uniquely produced by the prostate gland and elevated circulating concentrations are suggestive of increased prostate cell activity, which is sometimes attributable to prostate tumours and therefore, PSA can be used as a testing tool for prostate cancer [21, 31].

It is estimated that a PSA concentration cut-off value of 4 ng/mL has a sensitivity of 20-40% and a specificity of 90% and the sensitivity for the detection of high-grade cancer is 50% [32]. This low sensitivity means that some men with PSA levels below the cut-off will have prostate cancer. Furthermore, PSA levels are not specific to any prostate tumour; between 23-67% of prostate cancers detected by PSA testing may not have ever become clinically apparent [20, 22, 33].

The European Randomized Study of Prostate Cancer Screening Trial (ERSPC) estimated a 21% relative risk reduction in prostate cancer mortality after 13 years of follow-up [34], corresponding to one less death from prostate cancer per 1,000 men

screened. However, reductions in prostate cancer mortality were not statistically significant in either the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) [35, 36] or the Cluster Randomized Trial of PSA Testing for Prostate Cancer (CAP) trials [37]. Although there were high PSA testing attendance in the control arms of these trials and few prostate cancer deaths were observed in both trial arms [36, 37]. Differences in the implementation and setting between studies may account for the discrepancies in risk reduction [38]. All trials observed an increase in the "over-detection" of prostate tumours in the intervention arms (i,e, Gleason grade $\leqslant$ 6)[34˘37].

Despite the controversy surrounding PSA tests, it is commonly used as a screening tool for prostate cancer and a decision aid for clinicians for biopsy [39]. There is variation in the use of PSA testing by country. In the US, ∼44% of men aged 60-74 reported receiving a PSA test in 2013 [40], although PSA testing is now declining following a recommendation against the use of PSA-based screening for prostate cancer by the U.S. Preventive Services Task Force (USPSTF) [24, 40]. In the UK, cross-sectional studies estimate an annual testing rate of ∼6% in middle aged and older men [41, 42] and a longitudinal study reported that 39% of men aged 45-69 years had a PSA test over a ten-year period [43]. Men who undergo PSA testing are more likely to have higher levels of education and are more health conscious [43, 44].

From an epidemiological perspective, investigating potential prostate cancer risk factors is complex due to the high proportion of early asymptomatic and low-grade tumours which are detected by PSA tests, rather than due to the onset of clinical symptoms [20, 22, 33, 36]. Due to this "over-detection" of prostate tumours, a greater emphasis is placed on investigating potential risk factors for high-grade and/or advanced stage prostate tumours.

2.4 Classification of prostate cancer

The prostate cancer Tumour-Node-Metastasis (TNM) Classification of Malignant Tumours is a globally recognised cancer staging system devised in the 1970s [45] (Table 2.1). This staging system describes the spread of solid cancers throughout

the body. T describes the size and spread of the primary tumour to nearby tissues, N describes whether the tumour has spread to local lymph nodes, while M describes whether distant metastasis is present 2.1).

The Gleason score for prostate tumours originated from the results of a prospective cohort study in the 1960s-1970s which summarised the histological growth patterns of prostate tumours and compared the results to patient prognosis and staging [46, 47]. This system scores prostate tumours between 1-5 based on tumour differentiation. Gleason 1 represents a well differentiated tumour sample which is correlated with a favourable patient prognosis, whereas Gleason 5 indicates very low levels of tumour differentiation and a poor patient prognosis [46, 47]. The total Gleason score is calculated as the sum of the two Gleason scores, with the first half of the score based on the most common cell morphology with the second half based on the non-dominant cell pattern with the highest grade, therefore the maximum total Gleason score is ten [48].

In the American Joint Committee on Cancer (AJCC) system, the TMN stage, Gleason score and circulating PSA concentrations data are combined to classify prostate cancer staging at diagnosis as standard medical practice (Table 2.2).

Table 2.1: TNM tumour staging system for prostate cancer.

Category	TNM	Description
Localised disease	Tx	Primary tumour cannot be assessed
	T0	No evidence of primary tumour
	T1	Clinically inapparent tumor that is not palpable
	T1a	Tumour incidental histologic finding in 5% or less of tissue resected
	T1b	Tumour incidental histologic finding in more than 5% of tissue resected
	T1c	Tumour identified by needle biopsy found in one or both sides, but not palpable
	T2	Tumour is palpable and confined within prostate
	T2a	Tumour involves one-half of one side or less
	T2b	Tumour involves more than one-half of one side but not both sides
	T2c	Tumour involves both sides
Local extension	T3	Extraprostatic tumor that is not fixed or does not invade adjacent structures
	T3a	Extracapsular extension (unilateral or bilateral)
	T3b	Tumour invades seminal vesicle(s)
	T4	Tumour is fixed or invades adjacent structures other than seminal vesicles, such as external sphincter, rectum, bladder, levator muscles, and/or pelvic wall
Metastatic disease	N0	No positive regional lymph nodes
	N1	Metastasis in regional lymph node(s)
	M0	No distant metastasis
	M1	Distant metastasis

Data source: AJCC 7th Edition [49].
Abbreviations: AJCC=American Joint Committee on Cancer; TNM=tumour-node-metastasis.

Table 2.2: AJCC Prostate cancer staging.

Stage	T	N	M	PSA (ng/mL)	Gleason
Stage I	T1a-c	N0	M0	< 10	≤ 6
	T2a	N0	M0	< 10	≤ 6
	T1-2a	N0	M0	X	X
Stage IIA	T1a-c	N0	M0	< 20	7
	T1a-c	N0	M0	$\geq 10 < 20$	≤ 6
	T2a	N0	M0	< 20	≤ 7
	T2b	N0	M0	< 20	≤ 7
	T2b	N0	M0	X	X
Stage IIB	T2c	N0	M0	Any	Any
	T1-2	N0	M0	≥ 20	Any
	T1-2	N0	M0	Any	≥ 8
Stage III	T3a-b	N0	M0	Any	Any
Stage IV	T4	N0	M0	Any	Any
	Any	N1	M0	Any	Any
	Any	Any	M1	Any	Any

Data source: AJCC 7$^{\text{th}}$ Edition [49].
Abbreviations: AJCC=American Joint Committee on Cancer; M=metastasis; N=node; PSA=prostate-specific antigen; T=tumour.

In order to distinguish low risk prostate cancer from more clinically severe forms, tumours are often characterised by a combination of stage and Gleason grade. Analysis in this thesis will use the following commonly applied tumour categories:

Stage:

- Localised- TNM $\leq$T2 with no reported lymph node involvement or metastases (or AJCC stage I–II).
- Advanced- TNM $\geq$T3 and/or N1+ and/or M1 (or AJCC stage III–IV).

Gleason grade:

- Low-grade- Gleason score $\leq$7, or well/moderately/poorly differentiated.
- High-grade- Gleason score $\geq$8, or undifferentiated.

Aggressiveness:

- Not aggressive- TNM stage $\leq$T3 with no reported lymph node involvement or metastases.
- Aggressive- TNM stage T4 and/or N1+ and/or M1 and/or AJCC stage IV disease or death from prostate cancer.

Generally, after a high PSA test and/or following a positive digital rectal examination, patients will undergo a prostate biopsy. Over time the number of biopsy cores taken has increased, which is associated with greater rates of detection [50] and higher Gleason scores [51]. Consequently, the Gleason grading system recommendations for treatment have been updated to reduce disease over-treatment [52, 53].

2.5 Established risk factors

Age, family history of prostate cancer, genetic variants and racial/ethnic group are the only well-established risk factors for prostate cancer.

2.5.1 Age

Age is the strongest predictor of prostate cancer risk. Only 1.6% of all prostate cancer cases are diagnosed under the age of 50, incidence increases with age thereafter until the age of about 70 years (Figure 2.5). Autopsy studies indicate that approximately 50% of men between 70 and 80 years of age will show histological evidence of a prostate malignancy [54]. Therefore, the observed lower incidence of prostate cancer in men over the age of 75 years may be due to co-morbidities and lower PSA test attendance [55]. Older men are also more likely to be diagnosed with a more advanced stage prostate cancer and have lower rates of survival [30, 55].

The prostate is hypothesised to be a site of multiple neoplastic transformation events, many of which give rise to latent prostate cancer, but a large proportion of these latent tumours do not undergo critical activating events, which will cause the tumour to develop into a clinically meaningful disease [56]. The increase in prostate cancer risk with age may be due to a greater cumulative exposure to risk factors including i) oxidative/ DNA damage ii) telomere shortening [57] iii) inflammation [58] iv) endocrinological changes [59] and v) changes in gene expression with ageing [60]. The cumulative exposure to these activating events with ageing will increase the probability of tumour development and/or progression [56].

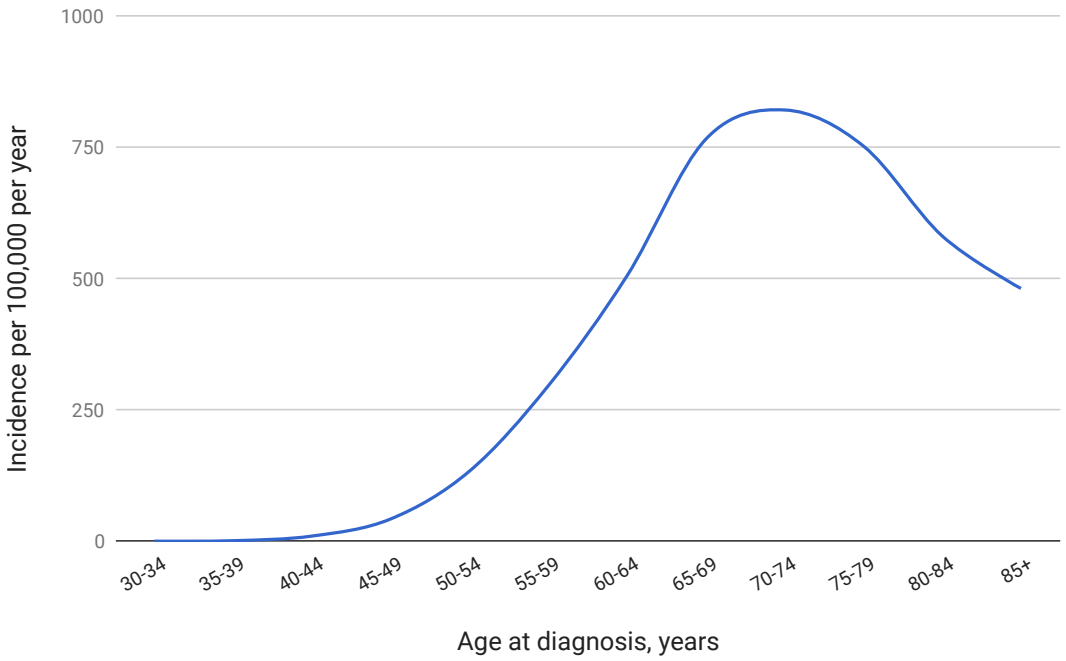


Figure 2.5: Prostate cancer incidence rates per 100,000 per year in US men by age at diagnosis (2005-2014).
Data source: [30].

2.5.2 Family history of prostate cancer

A meta-analysis found that having a first degree family history (brother or father) is associated with an increased risk of prostate cancer (relative risk (RR)=2.48, 95% confidence intervals (CI) 2.25-2.74), in comparison with men without a family history of prostate cancer [61]. The Nordic Twin Study estimated that genetic variants accounted for 57% of incident prostate cancer, making it one of the most heritable cancers [62]. However, it is probable that these estimates are inflated as individuals are more likely to be undergo a PSA test if a relative is diagnosed with prostate cancer [44, 63].

2.5.3 Genetics

The heritability of prostate cancer is determined by rare high-penetrant variants and more common variants which confer a lower risk [64]. Genome-wide association studies (GWAS) have so-far identified 167 single nucleotide polymorphisms (SNPs)

that are associated with prostate cancer [64, 65]. Currently identified prostate cancer susceptibility variants are estimated to account for ∼34% of familial relative risk of prostate cancer [64].

The highest penetrance mutation known for prostate cancer is the G84E mutation in the *HOXB13* gene, which is estimated to confer a 3.6-8.6 fold increased risk of overall prostate cancer [66]. This gene codes for a homeodomain transcription factor, which interacts with the DNA binding domain of the androgen receptor, altering its genomic recruitment sites [67]. Some of the other SNPs which are more modestly associated with prostate cancer risk have been mapped to the functionality of a number of transcription factors, some of which interact with the androgen receptor and/or androgen receptor co-factors as well as other non-androgen receptor pathways [68, 69].

Comparatively little is known about the genetic variants which may affect risk of advanced or aggressive prostate cancer. This may be due to the relatively small proportion of men with these tumour subtypes in each study [70].

2.5.4 Racial/ethnic group

As stated in Section 2.2.1, the incidence of prostate cancer is highest in high income countries and in countries with a large black population, while men in Asia have the lowest prostate cancer incidence [1, 2].

In the US, African Americans are estimated to have a 70% higher rate of diagnosis in comparison with non-Hispanic whites [71], despite comparatively lower PSA testing attendance [72]. This increased risk is also seen in the UK, where black men are estimated to have a two-to-threefold increased risk of being diagnosed with prostate cancer in comparison with whites [73, 74]. Furthermore, prostate tumours in African Americans generally occur in younger men, and present as a higher grade, and these men have a worse prognosis than non-Hispanic whites [30, 75, 76].

In the US, Hispanics have slightly lower rates of prostate cancer diagnosis than non-Hispanic whites [77], and a slightly lower survival rate [78]. In both the US and the UK, Asian men have a lower risk of prostate cancer in comparison

with non-Hispanic whites [74, 79]. Asian American men also present with a less favourable prostate tumour subtypes at diagnosis [75, 80], but have favourable survival in comparison with non-Hispanic whites [80].

The underlying causes of this ethnic/racial variation are not fully understood, but may have a substantial genetic component [72]. GWASs have identified susceptibility alleles which are more prevalent in African Americans [81–83]. This genetic variation may also lead to differences in enzyme and biomarker expression [84], including enzymes affecting sex steroid metabolism and circulating sex hormone concentrations [84–86]. Differences in diet and lifestyle may also play a role. For instance, Japanese-American men have a higher risk of prostate cancer than native Japanese men, this may be due to the adoption of a "Westernised" diet [87] and Asian immigrants in the US had a higher risk of prostate cancer than men in their native homeland even before the widespread use of PSA testing [88, 89].

The causes of these racial/ethnic disparities in prostate cancer risk are multi-factorial and complex; differences in PSA test attendance and treatment as well as genetics, hormones, lifestyle and cultural differences are likely to be involved in this disparity and render it difficult to establish any causal mechanism.

2.6 Putative risk factors

In 2014, the World Cancer Research Fund (WCRF) and American Institute for Cancer Research published "Diet, nutrition, physical activity and prostate cancer". This report contained a systematic review and meta-analysis of the evidence from prospective cohort and case-control studies and randomised controlled trials of prostate cancer as part of their aim to investigate how certain lifestyle factors may affect the risk of developing the disease [90]. This report concluded that height had a probable association with overall prostate cancer risk, while body fat had a probable association with advanced prostate cancer. Evidence of associations with physical activity, dairy consumption, high calcium diets, plasma selenium, and α-tocopherol with prostate cancer risk were more limited [90].

The following section summarises the evidence from the WCRF report and also includes the putative risk factors folic acid/folate, vitamin B_{12}, vitamin D, iron and phytoestrogens, which were not included in the original report due to the insufficient available evidence, but more evidence is now available. Associations with benign prostatic hyperplasia (BPH) and type II diabetes, which were out of scope for the WCRF report, are also outlined below.

This section also summarises the evidence from two international consortia; the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG) and the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium. The EHNBPCCG is a pooled dataset of prospective studies with hormone and nutritional biomarkers and prostate cancer case control status, while the PRACTICAL consortium holds genetic and prostate cancer status data.

2.6.1 Height

Taller adult height is associated with an increased risk of many cancers, including prostate cancer [91, 92]. Maximal attained height is generally determined by a combination of genetics and environmental influences, particularly childhood health and nutrition, which may also play a role in subsequent cancer risk [93]. Growth

hormone (GH), IGF-I and estradiol also play a direct role in the regulation of bone growth [94], and circulating IGF-I concentrations in adulthood are positively associated with prostate cancer risk [14]. It has also been hypothesised that taller people may have an increased risk of cancer as they have more cells and therefore, an increased probability of cellular replication errors [95].

In contrast with observational studies, Mendelian randomization studies do not clearly support a causal association between genetically determined height in adulthood and prostate cancer risk [96, 97], although the PRACTICAL consortium reported evidence that genetically taller men may have a higher risk of prostate cancer-specific mortality in men with low-grade disease (odds ratio (OR) per 1 standard deviation (SD) increase in height score=1.13; 95% confidence interval (CI) 1.08, 1.20) [97]. Associations were not present with other tumour subtypes or overall prostate cancer. The disparity between prospective observational and Mendelian randomization studies may indicate that the observed association between height and prostate cancer risk is due to early-life environmental exposures which influence adult height (such as foetal development, early-life environment and health, hormones and diet), as opposed to the shared genetic contributions of maximal height and prostate cancer risk [98]. The heterogeneity in the risk estimates by prostate cancer mortality may also indicate that the genetic determinants of height or height determining phenotypes (such as IGF-I) may be involved in prostate cancer progression.

2.6.2 Body fat

The WCRF reported that epidemiological studies generally show a positive association between BMI and advanced prostate cancer, estimating an 8% increased risk per 5 kg/m^2 higher BMI (95% CI 1.04-1.12) [90]. This 'advanced' disease definition included various aggressive forms of the disease, but it was not possible to differentiate between stage, grade and prostate cancer death due to the small number of available studies with these data. A more recent analysis in the European Prospective Investigation into Cancer and Nutrition (EPIC) observed an increased

risk of high-grade prostate cancer in obese men, but a lower risk of low-grade and localised prostate tumours [92].

Mendelian randomization studies are not clearly supportive of an association between genetically elevated BMI and overall prostate cancer risk [97, 99]. The PRACTICAL consortium reported weak evidence of a lower risk of localised prostate cancer with higher BMI (odds ratio (OR) per 1 SD increase in BMI genetic score=0.98, 95% CI 0.96-1.00 p=0.05), but these men had an increased risk of prostate cancer mortality (OR per 1 SD increase in BMI genetic score=1.08, 95% CI 1.03-1.14) [97]. The Genetic Associations and Mechanisms in Oncology (GAME-ON) consortium did not find evidence of an association between adulthood BMI and overall or aggressive risk of prostate cancer. However, there was evidence of an association between a higher birth weight and aggressive prostate cancer risk (OR=1.63 per 1 SD increase in birth weight 95% CI: 1.03-2.57; P=0.04) [99].

The weak associations observed in Mendelian randomization studies may indicate that environmental factors such as diet and physical activity account for, to a greater extent, the associations reported in conventional prospective epidemiological studies. Alternatively, the association between birth weight and aggressive disease may implicate the importance of early life factors in aggressive disease subtypes. However, it is possible that in both the observational and Mendelian randomization studies the associations may partly explained by detection bias; obesity-related haemodilution reduces PSA values and digital rectal examination is less accurate in obese men [100, 101]. Alternatively, associations may be partly explained by a lower attendance of PSA testing in more obese men [44]. This reduced the likelihood of early detection may lead to an increase risk of being diagnosed at a more advanced stage of prostate cancer.

These associations of BMI with high-grade prostate cancer may also be causal; for example, obesity is characterised as a chronic low-grade proinflammatory state. Chronic inflammation and endocrine disruption has a large impact on normal physiology, and leads to disruption of cell signalling, which may increase the probability of uncontrolled cell growth [102–104]. Because of the inherent complexity

of obesity and the probable role of detection bias, it is difficult to identify specific candidate signalling pathways or hormones which may explain the associations.

2.6.3 Co-morbidities

Benign prostatic hyperplasia (BPH)

BPH describes an enlarged prostate gland, and is thought to affect ∼70% of men aged ≥70 years [105]. Autopsy data indicates that ∼83% of prostate tumours were concomitant with BPH [106]. A meta-analysis of prospective studies found that the presence of BPH was associated with an increased risk of prostate cancer (RR=2.93, 95% CI 1.88-4.56) [107]. However, the Prostate Cancer Prevention Trial (where all enrolled men underwent annual digital rectal exam and PSA tests and were offered an end of study biopsy) did not show evidence of an association between BPH and prostate cancer risk [108]. This may indicate that the associations reported in general population studies is due to detection bias.

Type II diabetes

Observational prospective studies consistently report that type II diabetes is associated with a 14-30% reduced risk of overall prostate cancer diagnosis [74, 109–112], but the association, if any, with high-grade tumours is less clear [79, 110, 112].

The association with overall prostate cancer may be driven by physiological changes in hormone concentrations that are associated with type II diabetes, including androgens, IGFs, C-peptide and insulin [113, 114]. Detection bias might also contribute to the observed association with prostate cancer diagnosis. Men with diabetes are also less likely to attend testing [115] or to have a biopsy after a positive PSA test [79]. Further evidence of this detection bias was reported by the Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trial, in which participants underwent prostate biopsy regardless of serum PSA concentrations, and which found no evidence of an association between type II diabetes and prostate cancer risk [116].

2.6.4 Inflammation and infection

It has been widely hypothesised that chronic inflammation and the over-activation of the immune response may increase the risk of the onset and progression of cancer [117, 118]. There is some indication that inflammation and infection may be involved in prostate cancer development. but the epidemiological evidence is inconsistent [119–121] and it is difficult to rule-out detection bias [121].

2.6.5 Diet

The epidemiological evidence that diet may affect prostate cancer risk is inconsistent. The WCRF report stated that the evidence that a higher consumption of dairy products increases the risk of prostate cancer is limited but suggestive, while the evidence of other dietary factors and prostate cancer risk was too limited to draw a conclusion [90], To-date there is no clear evidence that any dietary factor may affect prostate cancer risk.

Dairy products

The WCRF concluded that the evidence that a higher consumption of dairy products increases prostate cancer is limited. This report estimated that risk of prostate cancer increases 7% per 400g daily consumption of total dairy products (RR=1.07, 95% CI 1.02-1.12), but there was insufficient power to stratify the associations by prostate cancer subtypes [90]. A more recent literature review also concluded that a higher consumption of milk products may increase the risk of prostate cancer progression [122]. Dairy products are high in calcium and higher dairy consumption is associated with higher circulating IGF-I concentrations [123]. Both calcium and IGF-I have been hypothesised to explain the observed association with prostate cancer [124, 125]. High calcium intakes may lower circulating 1,25-dihydroxyvitamin D concentrations [126], the most active form of vitamin D, which will inhibit the proliferation of prostate tumour cells within *in vivo* and animal models [127]. A recent Mendelian randomization study did not find evidence of a causal association between serum

calcium and prostate cancer [128]. This suggests that calcium does not explain a possible association between dairy consumption and prostate cancer risk.

The potential role of IGF-I in conferring prostate cancer risk is discussed in greater detail below (Section 2.7.4).

In summary, men who consume high amounts of dairy may have an increased risk of prostate cancer, but any underlying mechanism is unclear.

Selenium

Selenium is a mineral mainly found in seafoods, organ meat and, to a lesser extent, grains and dairy products (as well as in Brazil nuts, which are not widely consumed) [129].

The EHNBPCCG reported that nail, but not blood selenium concentrations were inversely associated with risk of total prostate cancer (OR per 80 percentile increase=0.29, 95% CI 0.22-0.40) and the association was observed in both non-aggressive and aggressive disease [130]. However, the Selenium and Vitamin E Cancer Prevention Trial (SELECT) and a Mendelian randomization study do not support the role of selenium in overall prostate cancer risk, and instead showed some evidence of a positive association with advanced prostate cancer [128, 131]. This may be related to the timing/duration of the exposure or may indicate that the observed association between selenium and prostate cancer diagnosis in prospective observational epidemiological studies may be due to residual confounding.

Folate, folic acid and vitamin B_{12}

Folate is a naturally occurring B vitamin found in various fruits and vegetables and a range of other foods, whilst folic acid is the synthetic form of folate which is more stable, and therefore is used in dietary supplements and for fortification of cereals [132]. There is limited research on their separate biological roles and so they are referred to as folates throughout this thesis. Vitamin B_{12} is found in meat, fish, eggs and milk.

Folates and vitamin B_{12} are important for normal DNA methylation, repair and synthesis [133, 134]. Methylation regulates gene transcription and cell signalling

[135]. Therefore, folate and vitamin B_{12} deficiencies may cause coding errors and double strand breaks in DNA replication and transcription [136], which may increase the risk of a carcinogenic event.

The EHNBPCCG reported that higher circulating folate and vitamin B_{12} concentrations were associated with an increased risk of overall prostate cancer incidence (OR highest fifth vs. lowest fifth=1.13, 95% CI 1.02-1.26 and OR=1.12, 95% CI 1.01-1.25, for folate and vitamin B_{12}, respectively). Higher folate concentrations were also associated with an increased risk of high-grade disease (OR highest fifth vs. lowest fifth=2.3, 95% CI=1.28-4.12) [137].

Randomised controlled trials do not support an association between folate supplementation and prostate cancer risk [138] and the genetic evidence of an association between folate and vitamin B_{12} and prostate cancer risk is inconsistent [139–141]. It has been suggested that the inconsistencies in the observed associations with prostate cancer risk may be due to variation of folate status between study populations [139].

Overall, it is possible that folate and/or vitamin B_{12} may be associated with prostate cancer risk, but further research is needed before it is possible to infer causality.

Vitamin D

Vitamin D is obtained from the synthesis of cholecalciferol from cholesterol through a chemical reaction that is dependent on sun exposure and can also be found in oily fish, meat, dairy foods, and some fortified foods. Vitamin D is important in cell growth, differentiation and apoptosis [142]. Polymorphisms in the vitamin D receptor gene may be associated with prostate cancer [143]. Circulating 25-hydroxyvitamin D (25(OH)D) concentration is regarded as the most informative indicator of vitamin D status [144], but some studies have also measured 1,25-dihydroxyvitamin D ($1,25(OH)_2D$), which is mainly formed by hydroxylation of 25(OH)D in the kidneys.

The EHNBPCCG reported an increased risk of prostate cancer in men with higher concentrations of 25(OH)D (OR highest fifth vs. lowest fifth=1.22, 95% CI

1.13–1.31) [145]. However, the association was only present in non-aggressive disease (OR per 80 percentile increase=1.24, 95% CI 1.13–1.36). 1,25$(OH)_2$D concentration was not associated with risk for prostate cancer overall or by tumour characteristics [145]. This may suggest that the association with non-aggressive disease is due to detection bias. Furthermore, a Mendelian randomization study did not support an association between 25(OH)D and risk of overall or aggressive prostate tumours [146].

2.7 Sex hormones and insulin-like growth factors

The evidence outlined above indicates that although there are large differences prostate cancer risk between populations, there are no well-established modifiable risk factors for prostate cancer. Furthermore, while prostate cancer may be associated with sociodemographic, anthropometric factors and co-morbidities, the underlying biology conferring these possible associations with risk are not understood.

Sex steroid hormones and IGFs have been implicated in prostate cancer aetiology and/or progression. The association between circulating testosterone concentrations and prostate cancer is unclear [13, 19], while associations with IGF-I are more robust [14]. Importantly, circulating testosterone and IGF concentrations are potentially modifiable. Therefore, if these associations with prostate cancer are confirmed, these hormones may explain some of the observed associations of the other characteristics outlined above with prostate cancer risk.

The following section provides a brief overview of sex hormone and IGF regulation, bioavailability, their action in the prostate, and summarises the current evidence of these hormones' involvement with prostate cancer risk.

There is also some evidence of an association between circulating sex hormone binding globulin (SHBG) and estradiol concentrations with prostate cancer risk, although this evidence is more limited, and this is briefly summarised in Section 2.7.4.

2.7.1 Sex hormones

Biosynthesis

Sex steroid hormones are hormones which interact with androgen or estrogen receptors. They are synthesised from cholesterol (Figure 2.6) and are secreted by the gonads, and adrenal glands, or formed by the conversion from other sex steroids in the peripheral tissues [147].

The first androgens on the sex steroid pathway are dehydroepiandrosterone (DHEA) and androstenedione; these are testosterone precursors (the main testicular androgen) (Figure 2.6). Testosterone exerts a biological effect by binding to the androgen receptor. Testosterone can also be converted to dihydrotestosterone (DHT)

by the enzyme 5α-reductase [147]. DHT has a higher affinity with the androgen receptor and therefore is considered a more biologically active form of testosterone. 5α-reductase is expressed across many androgen-responsive tissues including the prostate, skin, and testes, and acts to enhance the effect of testosterone [147, 148]. Androgens are metabolised into androstanediol glucuronide (A-diol-G), which is excreted and can also be used as a marker of androgen metabolism [149, 150].

Testosterone is also a precursor to the estrogens. It is converted into estradiol in the peripheral tissues by the enzyme aromatase, which is primarily located in the fat cells [151]. Although estrogens are found in much higher concentrations in premenopausal women, in men they still play important roles in adiposity, bone growth, and skin and lipid metabolism [147].

The adrenal cortex secretes DHEA, predominantly in the form of its sulphate ester dehyrdroepiandrosterone sulphate (DHEAS). This sulphate stabilises DHEA, increasing its half-life from approximately 1-3 hours to ∼10-20 hours [152]. DHEAS is thought to principally exert its effects via its downstream conversion into androgens and estrogens [153] (Figure 2.6).

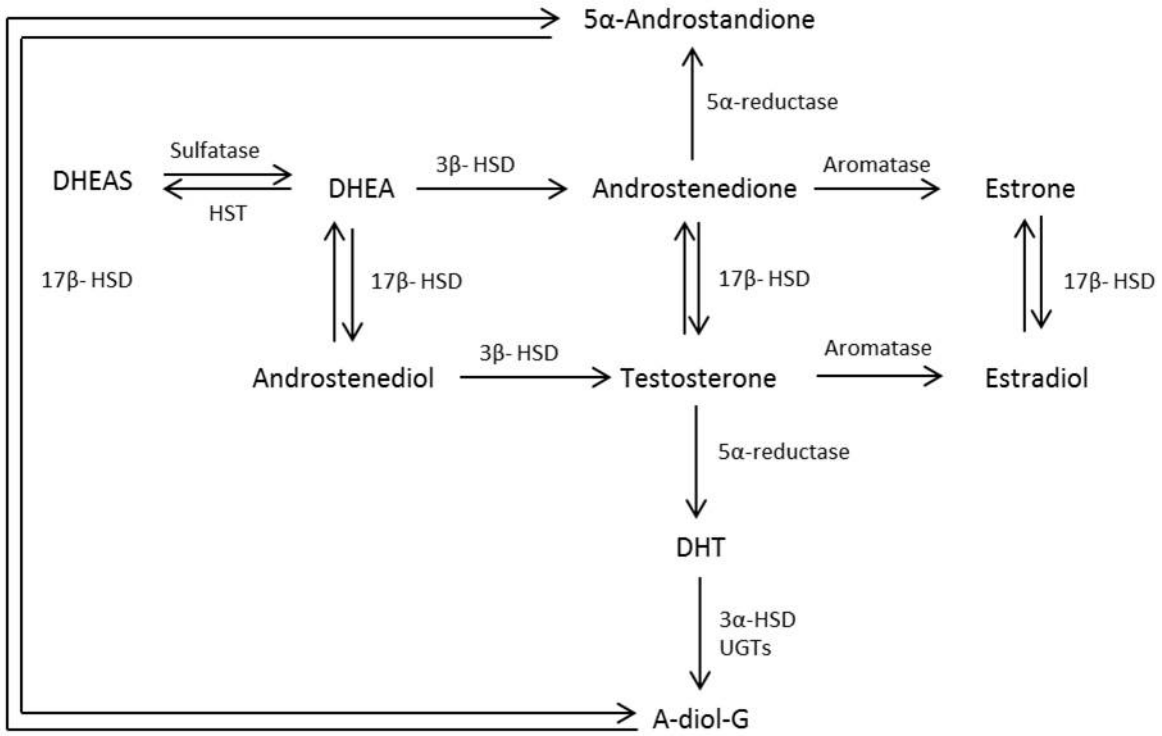


Figure 2.6: Simplified schematic of the biosynthesis pathway of sex hormones.
Abbreviations: A-diol-G=androstanediol glucuronide; DHEAS=dehydroepiandrosterone sulfate; DHT=dihydrotestosterone; HSD=hydroxysteroid dehydrogenase; HST=hydroxysteroid sulfotransferase; UGT=uridine 5'-diphospho-glucuronosyltransferase.

Steroidogenesis and regulation

The hypothalamic-pituitary-gonadal (HPG) axis plays a critical role in regulating testosterone production (Figure 2.7). The testes are the primary source of testosterone, producing approximately 95% of circulating testosterone [154].

The hypothalamus produces gonadotropin releasing hormone (GnRH) which is transported to the pituitary gland where it stimulates the synthesis of the gonadotrophic hormones luteinising hormone (LH) and follicle stimulating hormone (FSH). LH circulates in the bloodstream and is able to enter Leydig cells in the testes, where it stimulates testosterone secretion into the circulation [155]. Circulating testosterone can bind to androgen receptors in the hypothalamic neurons and in the pituitary gland, inhibiting LH production. FSH also stimulates the secretion of inhibin, which in turn inhibits FSH synthesis and secretion. This negative feedback loop enables testosterone to self-regulate its production [154].

The regulation of adrenal hormone steroidogenesis is less well characterised. Between 75-90% of DHEA is produced in the adrenal cortex, the remainder is produced by the testes [156]. Synthesis of adrenal DHEA is controlled by the hypothalamo-pituitary-adrenal axis. The hypothalamus produces corticotrophin releasing hormone (CRH) which stimulates the pituitary gland to synthesise adrencocorticotrophic hormone (ACTH), which stimulates the synthesis of DHEA and DHEAS by the adrenal gland. ACTH also stimulates the production of cortisol which inhibits CRH and ACTH synthesis, therefore DHEA(S) does not directly regulate its own production [156].

Sex hormones in circulation

Sex steroid hormones are lipid soluble and therefore when they are synthesised, they are able to pass out of the cell through the lipid membrane. A large proportion of sex hormones circulate in the blood bound to proteins. These proteins are unable to move out of the bloodstream into the peripheral tissues, and so store these hormones in the blood as an inactive bound component. These proteins also regulate the non-protein bound or “free” fraction of steroid hormones in the circulation.

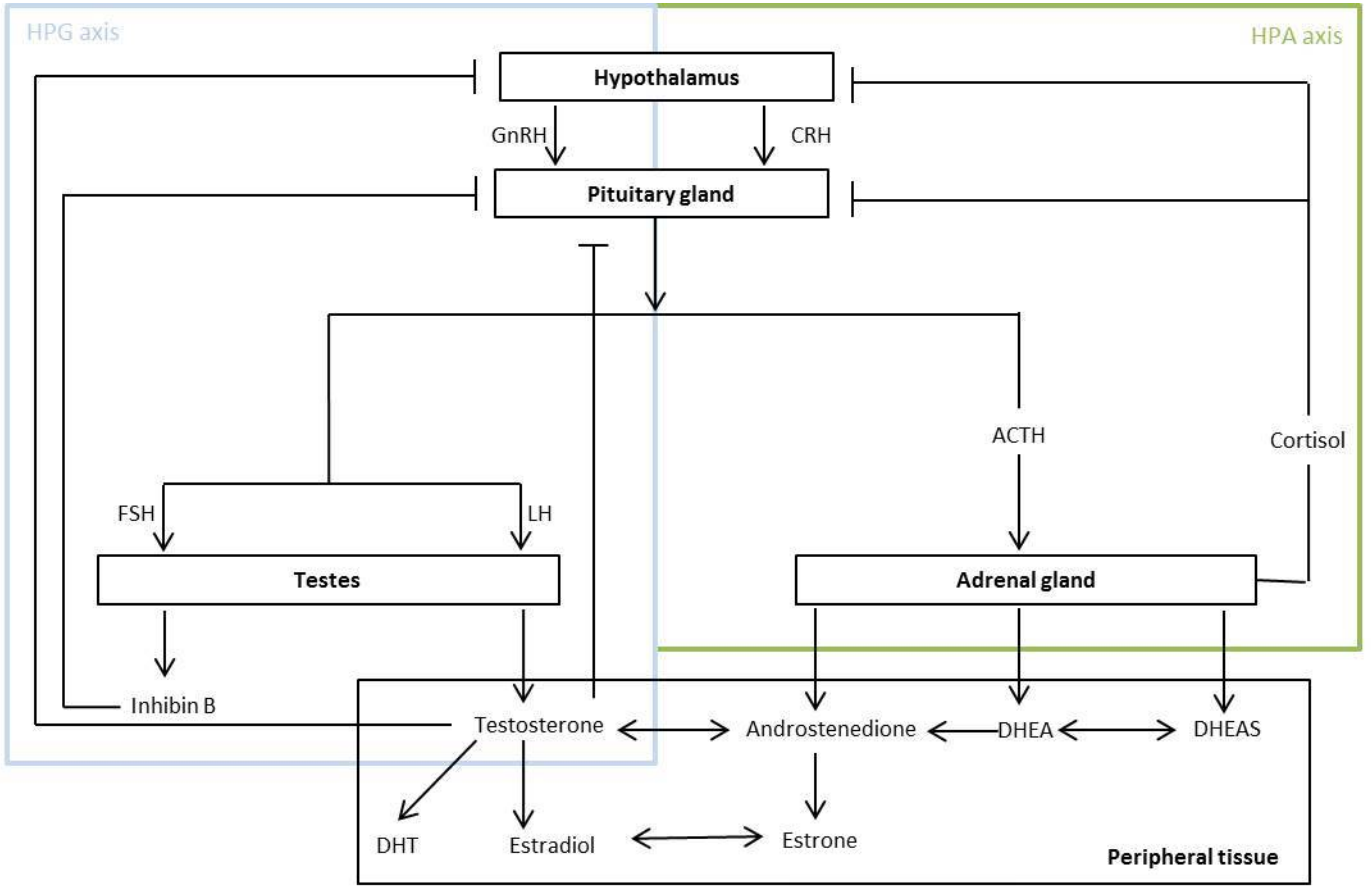


Figure 2.7: Simplified schematic for the main secretory agents and regulatory mechanisms of sex steroid hormones.
Abbreviations: ACTH=adrenocorticotropic hormone; CRH=corticotropin-releasing hormone; DHEA=dehydroepiandrosterone; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; FSH=follicle stimulating hormone; GnRH=gonadotropin-releasing hormone; HPA=hypothalamic pituitary adrenal; HPG=hypothalamic pituitary gonadal; LH=luteinizing hormone.

The free hormone hypothesis states that the biological activity of a hormone is determined by its free, rather than total, concentration [157]. The sex steroid hormones and binding proteins are at chemical equilibrium:

$$[Free\ sex\ steroid\ hormone] + [Carrier\ protein] \rightleftharpoons [Bound\ sex\ steroid\ hormone]$$

This provides a constant source of both active and "stored" hormone. If this equilibrium becomes unbalanced, then the components will rearrange to restore equilibrium. Despite the widespread adoption of the free hormone hypothesis, its validity remains the subject of scientific debate [158, 159].

Testosterone in circulation

Approximately 98% of circulating testosterone circulates bound to carrier proteins, of which SHBG is the most prevalent (bound to ∼50% of total testosterone [160]), followed by albumin and cortisol-binding globulin. In accordance with the free hormone hypothesis, testosterone that is bound to these carrier proteins is unable to exit the blood capillaries, therefore is considered biologically inactive. The remaining "free" 1-2% can pass out of the capillaries into target tissues, including the prostate where it is able to exert a biological effect.

SHBG is a glycoprotein which is predominantly synthesised in the liver. Its synthesis is stimulated by thyroid hormones and inhibited by IGFs, insulin and hepatic lipids [161, 162]. Its primary function is assumed to be that of a carrier protein. SHBG has a higher binding affinity to testosterone than to estradiol, and approximately 43-45% of testosterone and 30% of estradiol circulates bound to SHBG [160, 163].

SHBG also plays a role in the regulation of circulating testosterone concentrations. Increases in SHBG concentrations will reduce free testosterone concentrations by reducing the negative feedback of GnRH [164]. This increases LH secretion by the

pituitary gland, increasing testosterone production, thereby maintaining constant free testosterone concentrations [164].

Albumin is the most abundant protein in human circulation, accounting for ∼60% of the total serum protein content. Sex steroid hormones also bind to albumin, but with a lower affinity than to SHBG. Approximately 50% of circulating testosterone is bound to albumin [163, 165].

Total testosterone refers to the sum of the concentrations of both bound and unbound testosterone in circulation. Free testosterone is the sum of circulating testosterone which is not bound to any plasma protein.

2.7.2 Insulin-like growth factors

The IGF axis is a group of proteins consisting of the two major ligands (IGF-I and IGF-II), and a family of six binding proteins (IGFBP-1, IGFBP-2, IGFBP-3, IGFBP-4, IGFBP-5, IGFBP-6) [166, 167]. IGF-I and IGF-II are hormones with a high sequence similarity to insulin.

The IGFs exert their effects primarily through the type 1 IGF receptor (IGF-1R). IGF-II can also bind to the type 2 IGF receptor (IGF-2R), which channels the ligand for degradation within the cell, therefore it acts to inhibit IGF-II activity. IGF-1Rs can be activated by physiologic concentrations of IGF-I and IGF-II, but have a lower affinity with insulin [168]. IGFs are also able to bind to insulin receptors at a much lower affinity; IGF-II has a binding affinity to the insulin receptor, which is approximately 1/10 that of insulin, and IGF-I is 1/50 to 1/100 relative to insulin [169, 170]. IGFs can also exert their effects through hybrid insulin-IGF-I receptors [171].

The binding of IGFs to their receptors activates a number of signalling pathways, which are fundamental to cell growth, differentiation and survival.

The somatotropic axis

The IGF regulatory system is complex and has multiple determinants [166]. Approximately 85% of serum IGF-I is produced in the liver [172], but most tissues also have the ability to produce some IGF-I locally. At the most rudimentary level, circulating IGF-I synthesis is controlled by the somatotropic axis [173]. The hypothalamus secrets growth hormone-releasing hormone (GHRH), which stimulates the pituitary gland to secret growth hormone (GH). GH circulates in the blood stream and acts directly on the liver and other target tissues to stimulate IGF-I production [173]. IGF-I inhibits GH and GHRH production. This negative feedback loop enables IGF-I to self-regulate its production [173] (Figure 2.8).

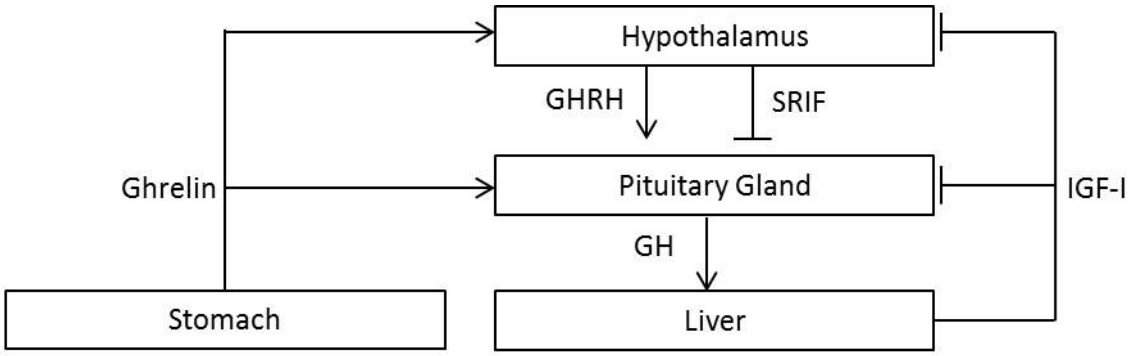


Figure 2.8: Somatotropic axis (simplified).
GH=Growth hormone; GHRH=Growth hormone-releasing hormone; IGF-I=Insulin-like growth factor-I; SRIF=Somatotropin release-inhibiting factor.

IGFs in circulation

The IGFs are not stored, instead they are constitutively secreted from the liver and in smaller amounts from other tissues and circulate at concentrations 1,000 times higher than most other peptide hormones. This high circulating concentration is maintained by IGFBPs. The majority of IGFs (75-80%) circulate bound to IGFBP-3 and acid-labile subunit (ALS), 20-25% of the bound IGFs circulate in binary complexes with IGFBPs 1-6 [167].

IGFBPs have several principal functionalities: i) creating a circulatory reservoir of IGFs by increasing the half-life of the IGFs from a few minutes (free IGFs), to 20-30 minutes in the binary complexes (IGF and IGFBP), to >16 hours in the ternary complexes (IGF, IGFBP and ALS) [174]; ii) modulating IGF activity at the cellular level by restricting their access to IGF receptors, iii) enhancing IGF receptor binding by concentrating IGF locally and increasing the availability for IGF receptor binding; iv) exerting IGF independent effects [167, 175, 176]. IGF action is further controlled by IGFBP proteases, which cleave IGF from their binding proteins [177].

A small proportion of IGFs also circulate unbound to IGFBPs (or "free"). Because of the complexity of IGFBP bound IGF in determining the bioavailability of the IGFs, and modulating receptor binding [167] as well as the expense and difficulty in measuring free IGF, previous research has placed less emphasis on free IGFs in comparison with free testosterone or free estradiol.

For the purpose of this thesis, IGF bioactivity may be simplified such that, in general, increases in peripheral IGF concentrations increase IGF bioactivity, whilst increases in IGFBP concentrations may generally decrease IGF bioactivity [178, 179].

2.7.3 Hormone action in the prostate

Understanding how circulating testosterone and IGF-I concentrations relate to intra-prostatic concentrations, and therefore biological activity, is fundamental to hypothesising how peripheral hormone concentrations may be involved in prostate cancer.

Prostate gland structure

The prostate is divided into the peripheral, transitional and central zones, the majority of prostate cancers originate from the peripheral zone [180]. These zones contain both epithelial and stromal cells (Figure 2.9). The stromal cells are the connective tissue which surround and support the epithelial cells. The epithelial compartment consists of slow growing well-differentiated luminal cells and rapidly proliferating basal epithelial cells, which differentiate into secretory luminal cells. There are also a small number of neuroendocrine cells [181]. It is widely thought that most prostate tumours originate from the epithelial cells [182].

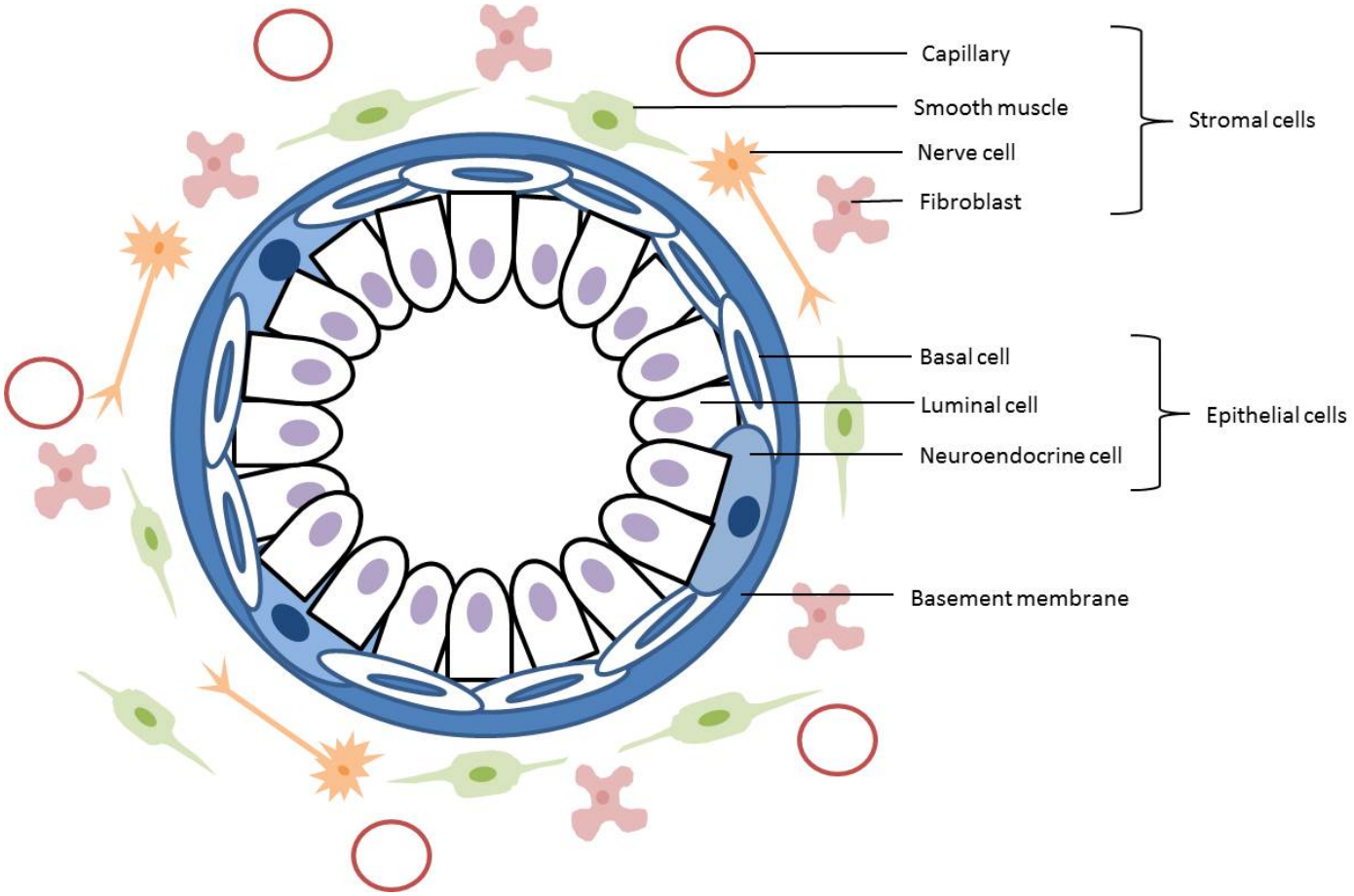


Figure 2.9: Prostate tissue structure.
Glandular epithelial cells consist of basal cells and luminal secretory cells. Fibromuscular stroma surrounds the basement membrane which consists of capillaries, smooth muscle cells, nerve cells and fibroblasts.

Intra-prostatic sex hormones

Androgen receptor signalling is integral for normal prostate development, the maintenance of tissue structure and cell differentiation [183]. Androgen receptors are expressed in both the stromal and luminal epithelial cells [184], and there is also a low expression of the androgen receptors in basal epithelial cells [185].

In the adult prostate, stromal androgen receptor signalling promotes epithelial cell growth and prostate regeneration. In the epithelial cells, androgen receptor activation inhibits basal cell proliferation and stimulates differentiated luminal cell survival [183, 186].

In rat models, decreasing DHT concentrations by castration reduces the number of secretory epithelial cells by 70% [187]. This reduction in epithelial cell numbers can be restored by exogenous androgen administration [188]. The androgen receptor also directly regulates PSA production [189].

DHT is a potent androgen receptor agonist and is found in high concentrations in prostatic tissue. However, circulating DHT cannot pass into prostate tissue as intra-prostatic DHT concentrations are greater than circulating concentrations [190, 191]. Peripheral free testosterone may therefore be the primary source of intra-prostatic androgen concentrations. However, studies that administer testosterone to eugonadal men have not reported corresponding increases in intra-prostatic DHT concentrations [192] or prostatic activity [193], although very low concentrations of circulating testosterone have been shown to affect prostate growth and biology [191, 194, 195].

There are two theories of why prostate tissue may be androgen dependent and yet may be relatively unresponsive to changes in circulating free testosterone concentrations which are described below: i) persistent intra-prostatic DHT; ii) the androgen saturation theory.

i) Persistent intra-prostatic DHT: The prostate gland has a high 5α-reductase expression. Free testosterone is able to pass out of the bloodstream and enter the prostate, where 5α-reductase catalyses the conversion of testosterone to DHT (Figure 2.10). This is known as the classical pathway.

The prostate may also be able to synthesise DHT independently of testosterone from 17-hydroxypregnenolone and 17-hydroxyprogesterone (Figure 2.10) [196]. This is known as the backdoor pathway. This backdoor pathway enables the prostate to synthesise some DHT independently of its supply of testosterone from the circulation. The backdoor pathway was first identified using animal models [197, 198]. Human prostates also express the enzymes which are involved in this backdoor pathway [199, 200]. A small randomised controlled trial (n=13) supports persistent intra-prostatic androgen concentrations. Men in this trial received a GnRH antagonist which led to a 94% reduction in circulating testosterone concentrations (mean testosterone concentration=0.9 ± 0.3 nmol/L). This resulted in a 37% decrease in prostatic androgen receptor expression, and a 17% decrease in PSA concentrations, although these changes were not statistically significant (possibly due to the low power) [191]. Given the dependence of the prostate on androgen receptor signalling, this may indicate persistence of androgen receptor signalling, despite the low supply of testosterone from the circulation. A laboratory study of prostate cancer cell lines, which inhibited the enzymes involved in both the classical and backdoor pathways, was able to decrease DHT concentrations in prostate cancer cells to a greater extent than those that only inhibited one pathway [201].

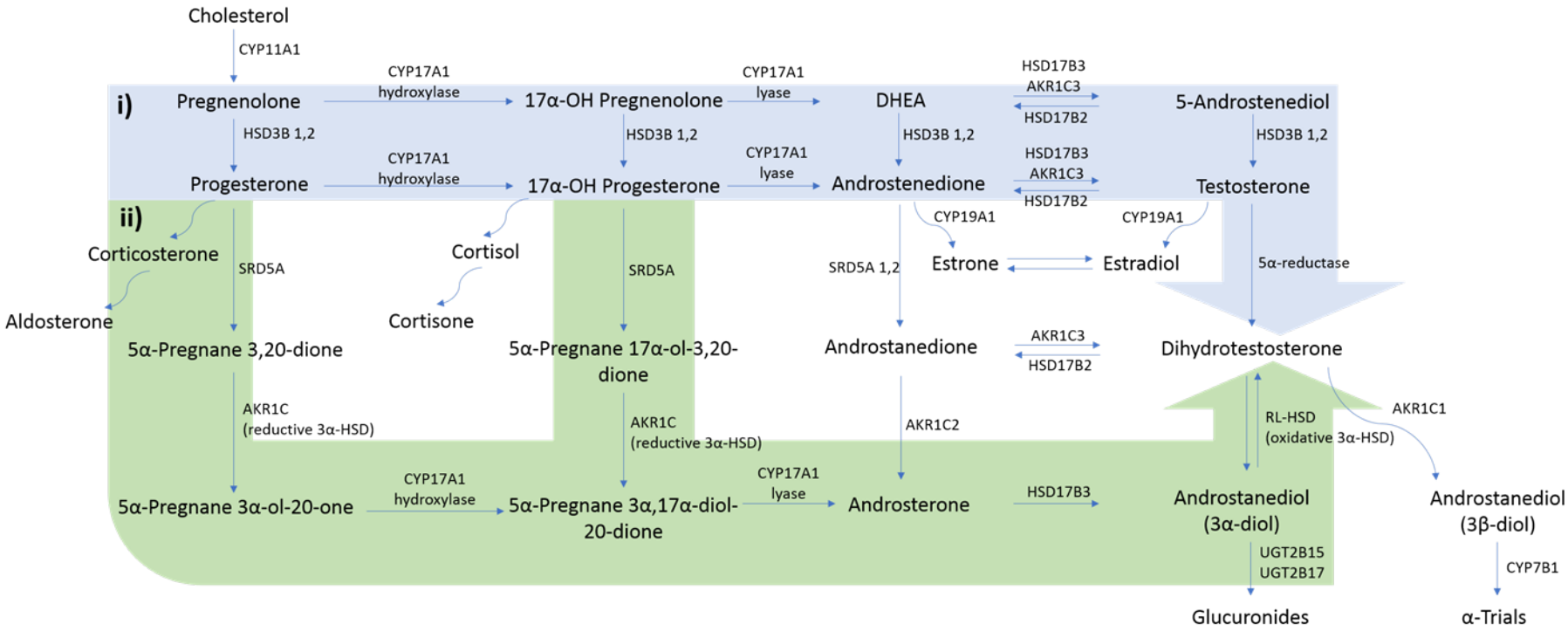


Figure 2.10: Schematic diagram of DHT synthesis in the prostate.
i) Classical pathway: testosterone is converted into dihydrotestosterone by 5α-reductase (blue arrow).
ii) Backdoor pathway: dihydrotestosterone is synthesised independently of testosterone from the progesterone precursors which are converted to androstanediol, this can be converted to dihydrotestosterone with an oxidative 3α-HSD enzyme (green arrow).

Abbeviations: AKR1C=Aldo-Keto Reductase Family 1 Member C; CYP=Cytochrome P450; DHEA=dehydroepiandrosterone; HSD=hydroxysteroid dehydrogenase.
Figure redrawn based on [196].

ii) The androgen saturation theory: The androgen saturation theory states that the stimulatory effects of androgens on the prostate will increase up to a certain concentration, above which the androgen receptors will become saturated and any additional increase in androgen concentrations will not induce a further biological effect [19] (Figure 2.11). The saturation of androgen receptors has been demonstrated in rat prostates [202], and within *in vitro* human fibroblast cell culture [203]. Furthermore, exogenous testosterone administration may stimulates an increase in PSA concentrations in hypogonadal, but not eugonadal men [204].

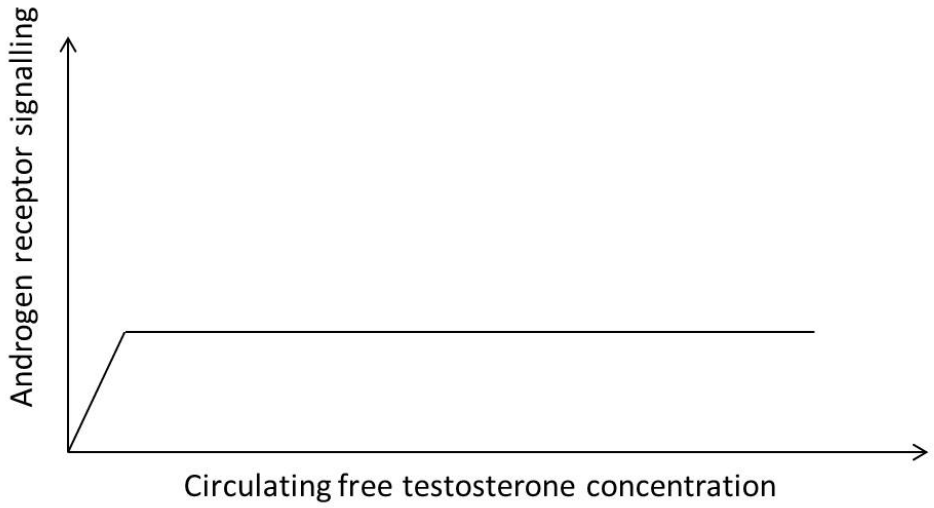


Figure 2.11: The androgen saturation model of intra-prostatic androgen signalling in relation to circulating free testosterone concentration.

Implications for prostate cancer risk

These two theories indicate that the prostate can maintain constant androgen receptor activity, independent of high-to-low circulating free testosterone concentrations. But when circulating free testosterone concentrations are very low this may reduce androgen receptor signalling (via decreases in intra-prostatic DHT and/or the "unsaturation" of the androgen receptors). This reduction may lead to changes in prostate growth and function and may therefore affect prostate cancer risk.

Intra-prostatic IGFs

Circulating IGF-I and IGF-II can pass out of the bloodstream and into prostate tissue and exert an effect via receptor binding. IGF-1R are expressed on the membrane of prostate epithelial cells and IGF-I is approximately ten-fold more potent than IGF-II in stimulating prostate epithelial cell proliferation, due to its greater affinity with the IGF-1R [205]. The binding of IGF ligands to IGF-1R activates two major signalling pathways; i) The phosphoinositide 3-kinase (PI3K)/Akt pathway, which induces a metabolic response and is also involved in cell growth and protein synthesis [206]; ii) the MAPK/ERK pathway which is involved in gene transcription and regulation [207] (Figure 2.12). The prostate is also able to produce small quantities of IGFs and IGFBPs locally [208].

Circulating IGF-I concentrations are thought to be broadly representative of prostate tissue IGF-I bioactivity, [209, 210], but the role of the IGFBPs in this relationship is less clear [167].

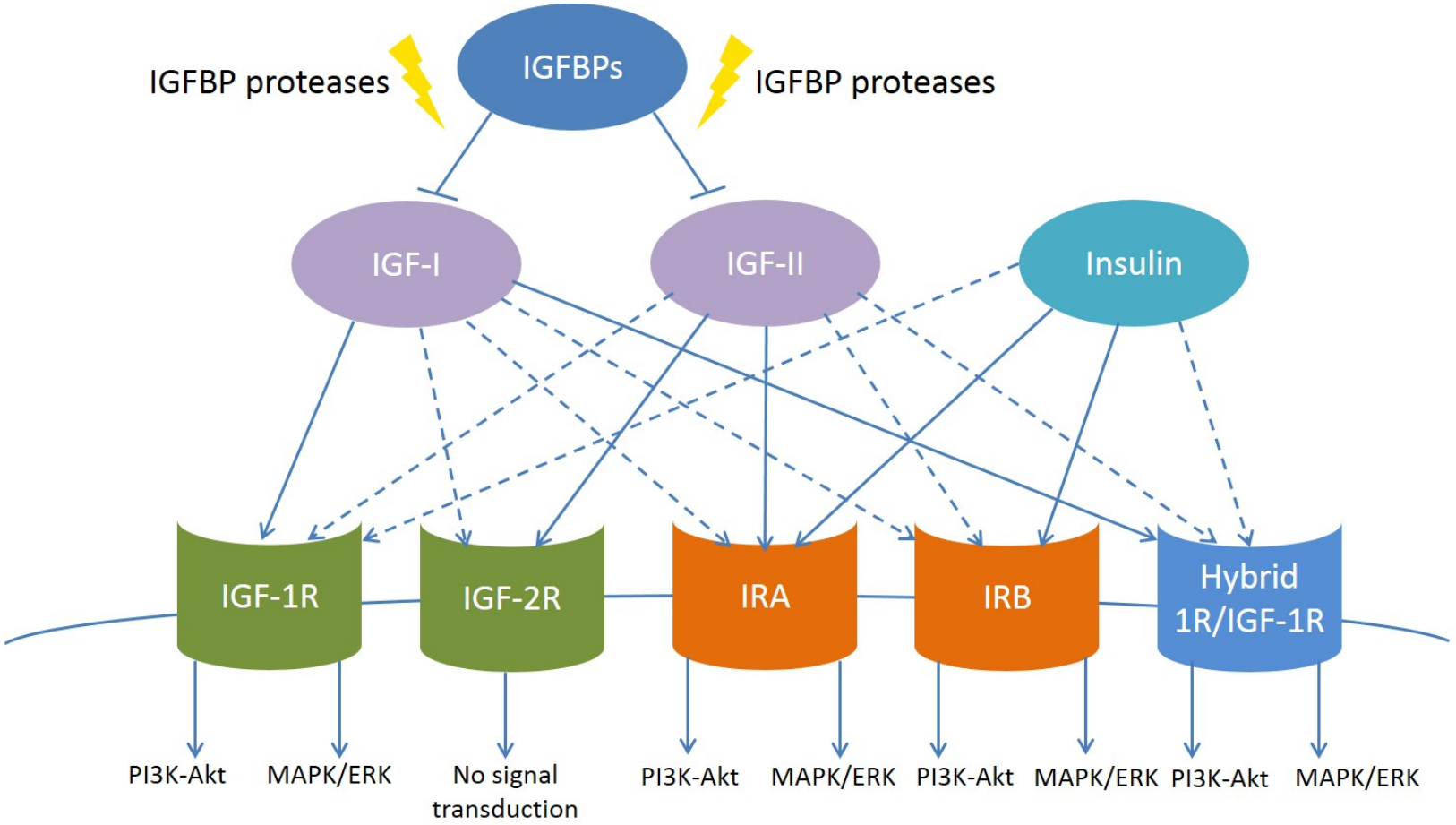


Figure 2.12: IGF signalling in the prostate (simplified).
Dashed arrow denotes a low receptor affinity.
IGF-I=Insulin-like growth factor-I, IGF-R=Insulin-like growth factor receptor, IGFBP=Insulin-like growth factor binding factor, IR=Insulin-receptor, PI3K-Akt=phosphoinositide 3-kinase-Akt, MAPK/ERK=Mitogen-activated protein kinase/extracellular signal-regulated kinases.

2.7.4 Hormones and prostate cancer

Testosterone

The androgen hypothesis states that prostate cancer development and progression is driven by androgens. This theory originates from work by the Nobel Prize winner Charles Huggins and Clarence Hodges who demonstrated that castration led to prostate cancer regression in men with metastatic disease, whilst exogenous testosterone administration led to tumour progression [12, 211]. This androgen dependence of prostate tumours may also indicate that testosterone is involved in prostate tumour initiation. Exogenous testosterone administration in rat models can induce prostate tumours [4, 5], and GWASs in humans have identified SNPs which are associated with prostate cancer risk in regions which code for proteins that interact with or are localised to the androgen receptor [6, 7].

The Prostate Cancer Prevention Trial and REDUCE trial are two randomised controlled trials which aimed to reduce prostate cancer incidence using 5α-reductase inhibitors, which will reduce intra-prostatic DHT concentrations. Both trials reported a 23–25% reduction in overall incident prostate cancer in men who received the intervention in comparison with controls [8–10]. However, both trials also reported a possible increased relative risk of high-grade prostate tumours [9, 10].

Despite this apparent strong biological and clinical evidence which supports the role of testosterone in prostate cancer risk, there is no prospective epidemiological evidence to support an association between circulating testosterone and prostate cancer risk. The EHNBPCCG has the largest pooled dataset of hormones and prostate cancer to-date. In 2008, this international collaboration analysed almost 4,000 prostate cancer cases from 18 prospective studies, and found no evidence of an association between any circulating sex hormone concentrations and prostate cancer risk [13]. The evidence outlined above suggests that there may be a threshold association between circulating free testosterone and prostate cancer risk. Therefore, previous prospective epidemiological studies may have lacked the power to test this possibly non-linear association. Since 2008 the number of recorded prostate cancer

cases in the EHNBPCCG dataset has approximately doubled. Thus, it may now be possible to test this possible threshold association with prostate cancer risk.

IGF-I

Prospective epidemiological studies indicate that circulating IGFs are associated with a number of cancers including prostate [14], colorectal [212], breast [213] and possibly thyroid cancer [214].

IGF-I is hypothesised to increase prostate cancer risk by increasing prostatic IGF-IR signalling, which increases the pool of damaged cells by reducing the programmed cell death, and this accumulation of damaged cells increasing the probability of carcinogenesis [215].

In animal models, overexpression or constitutive activation of IGF-1R promotes tumour formation [15]. *In vitro* studies of prostate tumour tissues show an increased IGF, IGFBP and IGF-1R expression [216] and a reduction in IGF-2R expression [217] (this reduction in IGF-2R expression will increase IGF-II bioactivity). Furthermore, men with tumours with high IGF-1R expression had an elevated risk of the tumour progressing to metastases and prostate cancer death [218]. Clinical studies show higher serum concentrations of IGF-I and IGFBP-2 in men with prostate cancer [16, 17].

The EHNBPCCG analysed 15 prospective studies including 7,682 prostate cancer cases with IGF-I measurements, and reported a positive linear association between circulating IGF-I concentrations with prostate cancer risk (OR highest fifth vs. lowest fifth=1.29, 95% CI 1.16-1.43) [14]. A Mendelian randomization study using the PRACTICAL consortium data (22,936 prostate cancer cases matched to 22,936 controls), also supported the possible involvement of the IGF pathway in prostate cancer progression and mortality, but due to pleiotropy of the genetic instruments with multiple IGF biomarkers, this analysis was unable to identify specific causal components [219].

The IGF system also plays a role in prostate tumour growth and progression. Increased IGF-1R and/or low PTEN signalling is associated with an increased risk

of progression to lethal disease [220]. PSA is a protease which cleaves IGFBP-3, and this proteolytic cleavage increases the bioavailability of IGF-I [221]. As prostate tumours develop, epithelial stromal IGF signalling becomes autocrine such that epithelial cells are able to grow independently of stroma cells [222].

Estrogen

In animal models, aromatase knockout rats (unable to synthesise estrogen) do not develop prostate cancer [223]. Further, combined estrogen and testosterone administration accelerates the rate of tumorigenesis, in comparison with testosterone alone [224]. In humans, estrogen receptors are expressed in healthy prostate tissue. High-grade prostatic intraepithelial neoplasia (HGPIN) describes an abnormality of the prostate tissue, which is thought to precede the development of prostate cancer. In this premalignant state, estrogen receptor subtype expressions are altered, and high-grade tumours also overexpress estrogen receptors [225].

However, a clinical trial did not show a relative risk reduction in the development of prostate cancer in men with HGPIN treated with toremifene (a selective estrogen receptor modulator) [226]. Furthermore, the EHNBPCCG did not find evidence of an association between circulating estradiol and prostate cancer risk [13].

Sex hormone binding globulin

SHBG binds to circulating testosterone and estradiol, and is a major regulator of their respective bioavailabilities [227]. Insulin, cytokines and some dietary components, including monosaccharides and palmitate play a role in SHBG regulation [228].

The EHNBPCCG reported an inverse linear association between circulating SHBG concentrations and prostate cancer risk (RR highest fifth vs. lowest fifth=0.86, 95% CI 0.75-0.98). The primary function of SHBG is a binding protein. Therefore, it is likely that this association is due its correlation with another factor. After further adjustment for IGF-I, the association observed in the EHNBPCCG attenuated slightly, which may indicate that IGF-I, which is weakly positively correlated with SHBG, may be partly driving the association [13].

2.8 Conclusion

Prostate cancer is the second most common cancer in men globally and the fifth leading cause of cancer death. The observed heterogeneity between countries may be partly due to genetic factors and modifiable risk factors. Differences in diagnostic practices; primarily, the frequency and use of PSA testing are also likely to explain some of the reported global variation and increased diagnosis.

Laboratory and clinical evidence support a threshold association between circulating free testosterone and intra-prostatic androgen signalling. This model may also apply to prostate cancer, whereby men with very low free testosterone concentrations may have a reduced risk of prostate cancer. However, men with moderately low-to-high circulating free testosterone concentrations may not show further dose-response increases in incident prostate cancer. Previous prospective studies may have been underpowered to test this non-linear association. The positive linear association of circulating IGF-I concentrations with prostate cancer risk is more established.

Both testosterone and IGF-I are also associated with a number of possible prostate cancer risk factors including racial/ethnic group, body fat, type II diabetes and possibly dairy consumption. Therefore, these hormones may be modifiable and account for some of the observed associations with prostate cancer.

In order to enhance the understanding of whether sex hormones and IGFs may be modifiable, and therefore may influence disease risk, research in this thesis will examine the possible cross-sectional anthropometric, sociodemographic and behavioural associations with circulating concentrations of these hormones. Research in this thesis will also examine the hypothesised roles of free testosterone and IGF-I in prostate cancer risk, in the context of the shapes of the associations described above, using both measured free testosterone concentrations and the clinical surrogates of low testosterone and high IGF-I. To further examine of role of testosterone as well as other potential risk factors for prostate cancer this thesis will also investigate the associations of haematological parameters with prostate cancer risk.

3

Studies used in this thesis

Contents

3.1 Introduction 51
3.2 The Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group 51
3.2.1 Introduction 51
3.2.2 Key contributors to the EHNBPCCG dataset and analysis 52
3.2.3 Study identification 52
3.2.4 Participating studies 52
3.2.5 Ethical approval 60
3.3 English national health records 61
3.3.1 Introduction 61
3.3.2 Hospital Episode Statistics Admitted Patient Care (HES APC) 61
3.3.3 Ethical approval 63
3.4 UK Biobank 65
3.4.1 Introduction 65
3.4.2 Design and methods 65
3.4.3 Ethical approval 69
3.4.4 Follow-up 70

3.1 Introduction

Analyses in this thesis used three different datasets:

1. Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG): an international pooled consortium of prospective studies of biomarkers and prostate cancer risk.

2. English national Hospital Episode Statistics Admitted Patient Care (HES APC) data and mortality records: national hospital admissions which are linked to mortality records.

3. UK Biobank: a large prospective cohort study.

This chapter summarises the design and data collection methodologies of these three datasets.

3.2 The Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group

3.2.1 Introduction

The EHNBPCCG was established in 2004 with the aim of collating worldwide prospective data in order to perform individual participant meta-analyses of circulating hormones and nutritional biomarkers in relation to prostate cancer risk.

As lethal prostate cancer is relatively uncommon, most individual studies do not have the statistical power necessary to detect an association with more aggressive prostate tumour subtypes. Therefore, the pooling of individual-participant level data is necessary to robustly estimate associations. To-date over 40 studies worldwide have contributed to this collaborative dataset, providing data on a total of approximately 36,000 prostate cancer cases and 56,000 male controls for more than 50 different biomarkers.

The EHNBPCCG currently contains 29 studies with insulin-like growth factors (IGFs) and/or sex hormone measurements, which are used in analyses in this thesis.

For simplicity, this chapter only describes the studies contained in this dataset which are included analyses in Chapters 4-6.

3.2.2 Key contributors to the EHNBPCCG dataset and analysis

The EHNBPCCG dataset was created and stored by the Cancer Epidemiology Unit, University of Oxford. The principal investigators are Timothy Key and Naomi Allen, who led the concept and search strategy. To enable the centralised pooling of data, Paul Appleby harmonised the individual studies and also maintained the dataset. Naomi Allen, Aurora Perez-Cornago, Timothy Key and Ruth Travis co-ordinated and contacted studies to request data.

I conducted and led all analyses which used the EHNBPCCG dataset in this thesis.

3.2.3 Study identification

Principal investigators were invited to join this collaborative group if they had published or unpublished data on prostate cancer risk and endogenous hormones and/or nutritional biomarker concentrations that had been determined from blood samples collected prior to diagnosis.

Studies were identified by literature searches of computerised bibliographic systems, including PubMed, Web of Science, Cochrane Library, and CancerLit, and through discussions with colleagues [13, 14].

3.2.4 Participating studies

Individual participant data were available from 29 studies with circulating IGF and/or sex hormone measurement information: Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC) [229], Baltimore Longitudinal Study of Aging (BLSA) [230, 231], British United Provident Association (BUPA) [232], Carotene and Retinol Efficacy Trial (CARET) [233], Child Health and Development Studies (CHDS) [234], Cardiovascular Health Study (CHS) [235], the CLUE I "Give us

a Clue to Cancer" Study [236], European Prospective Investigation into Cancer and Nutrition (EPIC) [237, 238], the European Randomized Study of Screening for Prostate Cancer (ERSPC) [239], Finnish Mobile Clinic (FMC) [240], Health in Men Study (HIMS) [241, 242], Helsinki Heart Study (HHS) [243, 244], Health Professionals Follow-up Study (HPFS) [245], Japan Collaborative Cohort Study (JACC) [246], Japan-Hawaii Cancer Study (JHCS) [247, 248], Japan Public Health Center-based prospective study (JPHC) [249], Janus [244, 250], Kaiser Permanente Medical Care Program (KPMCP) [251], Melbourne Collaborative Cohort Study (MCCS) [252, 253], Multiethnic Cohort (MEC) [254], Massachusetts Male Aging Study (MMAS) [255], Northern Sweden Health and Disease Cohort (NSHDC) [244], Prostate Cancer Prevention Trail (PCPT) [10], Physicians' Health Survey (PHS) [256], Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) [257], Prostate Testing for Cancer and Treatment (ProtecT) feasibility [258, 259] and main study [260], Rancho Bernado Study (RBS) [261], and the SUpplémentation en VItamines et Minéraux AntioXydants (SU.VI.MAX) trial [262].

Principal investigators were asked to provide data on prostate cancer case or non-case status and, if applicable, matched-set identifier. Data were also supplied on participant characteristics including date and age at blood collection, height, weight, racial/ethnic group, smoking status, alcohol consumption, education, marital status, and tumour characteristics data including date and age at diagnosis and tumour stage and grade.

Study designs

The characteristics of each study included in the collaborative analyses are found in their original publications, are summarised below and in Table 3.1.

The majority of the studies are case-control studies nested within traditional prospective cohort studies (CHDS, CHS, CLUE, EPIC, FMC, HPFS, JACC, JHCS, JPHC, Janus, MEC and RBS), a health plan (BUPA and KPCMP), or randomised trials (ATBC, CARET, ERSPC, HHS, NSHDC, PCPT, PHS, PLCO, ProtecT feasibility, ProtecT and SU.VI.MAX) and the original matched case-control sets were

retained. Of the randomised trials, data were available for participants in both the intervention and placebo arm, with the exception of the PLCO, where blood samples were only collected in the screening arm [263] and in the placebo arm of the PCPT (as the intervention arm was designed to alter prostate hormone concentrations) [10].

The ERSPC, PCPT, PLCO, ProtecT (feasibility and main) are observational investigations using data from trials that included organised screening for prostate cancer. In the ERSPC, PCPT and PLCO trials, men with a raised prostate-specific antigen (PSA) or abnormal digital rectal examination at recruitment-screening were excluded, and the eligible cases were diagnosed during subsequent follow-up. In the ERSPC and PCPT, most cases were diagnosed at the end of the study. The ERSPC study closed 4 years after recruitment and cases were detected through PSA screening. In the PCPT, the study closed 7 years after recruitment and prostate tumours were detected through a routine end of study biopsy. In the PLCO, most cases were detected by routine PSA testing, which was organised at regular intervals throughout the trial. The ProtecT was designed to investigate different prostate cancer treatments, in which (mostly asymptomatic) men were screened with PSA at study baseline [259]. Men with PSA $\geq$3 ng/mL were offered a diagnostic biopsy. This meant that 50% of cases were diagnosed within 3 months of blood collection and 58% within the first year. Therefore, the data from ProtecT are considered to be cross-sectional and as such, were not included in the prospective analyses (Chapter 6).

Four studies were cohort or case-cohort analyses (BLSA, HIMS, MCCS and MMAS), in which hormone concentrations had been measured from stored serum from all, or a subset, of the cohort. For inclusion in this collaborative dataset, the cases from each case-cohort study were matched at random to up to three control subjects who were free of prostate cancer at the time of diagnosis of the case and met the standard matching criteria (age at recruitment, date of recruitment, time of blood collection and race). This conversion of cohort or case-cohort studies to nested case-control studies enabled all analyses to be conducted on matched sets but resulted in some loss of power for each individual study.

There was some variation in the case mix of these studies according to the prevalence of PSA testing within that population during follow-up. For example, there is a generally higher proportion of early stage and low-grade cases in studies based in the US, where there has been relatively high levels of PSA-testing since the early-1990s, than in studies in European populations, where PSA-testing has only become common more recently.

Table 3.1: Collaborating studies in the Endogenous Hormones, Nutritional Biomarkers Prostate Cancer Collaborative Group dataset.

Study	Cases/controls	Country	Year of blood collection	Age of blood collection	Median years from blood collection to diagnosis (IQR)	Study design
ATBC IGFs	100/311	Finland	1985-1988	51-70	9 (2)	Randomised trial of α-tocopherol and β-carotene among smokers
ATBC hormone	116/232	Finland	1985-1988	50-69	4 (3)	
BLSA IGFs	72/111	USA	1963-1993	43-83	9.5 (1.5)	Prospective cohort study of the physiology of aging
BLSA glucose	114/114	USA	1961-1995	30-84	17 (16)	
BUPA	141/423	UK	1975-1982	36-64	11 (5)	Prospective cohort study of health plan members
CARET	298/298	USA	1986-1996	47-77	2 (3)	Randomised controlled trial of β-carotene and retinyl palmitate in heavy smokers and asbestos workers
CHDS	322/650	USA	1959-1967	20-55	33 (6)	Prospective cohort study of the husbands of pregnant women who were health plan members
CHS	174/174	USA	1989-1993	65-89	3 (2)	Population-based prospective cohort study
CLUE I DHEA	80/101	USA	1974	44-87	7 (5)	
CLUE I IGFs	30/60	USA	1974	42-79	12 (4)	Population-based prospective cohort study
CLUE I horm	96/96	USA	1974	44-87	7 (4)	
EPIC I	977/1,079	Europe	1992-1999	43-76	3 (3)	
EPIC II	1,198/1,211	Europe	1992-1999	39-77	6 (3)	Population-based prospective cohort study
EPIC III/IV	1,787/1,787	Europe	1992-1999	36-79	11 (4)	
EPIC-Norfolk	335/1,225	UK	1993-1998	45-78	4.5 (5)	
ERSPC	201/201	Netherlands	1993-1988	54-69	4 (0)	Population-based randomized trial of PSA screening
FMC	166/300	Finland	1968-1972	17-80	13 (9)	Population-based prospective cohort study
HHS	84/291	Finland	1980-1982	40-56	11 (5)	Randomised controlled trial of gemfibrozil
HIMS	336/1,325	Australia	2001-2004	71-87	3 (4)	population-based randomized trial of screening for abdominal aortic aneurysms
HPFS I	682/682	USA	1993-1995	46-87	3 (3)	Cohort study of male health professionals
HPFS II	1,326/1,326	USA	1993-1995	46-81	7 (2)	

Continued on next page

Table 3.1 – continued from previous page

Study	Cases/controls	Country	Year of blood collection	Age of blood collection	Median years from blood collection to diagnosis (IQR)	Study design
JACC	40/101	Japan	1988-1992	57-85	5 (4)	Population-based prospective cohort study of participants attending screening programs
JHCS 1988	98/98	USA	1971-1975	52-74	12 (7)	Prospective cohort study of Japanese men in Hawaii
JHCS 1996	143/143	USA	1971-1975	52-74	12 (5)	
JPHC	201/402	Japan	1990-1994	41-71	10 (6)	Population-based prospective cohort study
Janus	641/2,498	Norway	1972-1990	33-66	17 (7)	Population-based prospective cohort study of participants in the Norwegian Regional Health Studies and Red Cross blood donors
KPMCP	45/218	USA	1964-1970	60-85	8 (7)	Prospective cohort study of health plan members
MCCS	591/1,056	Australia	1990-1994	40-72	5 (5)	Prospective cohort study of inhabitants of the Melbourne metropolitan area
MEC	467/934	USA	1994-2004	48-85	1 (2)	Ethnically diverse prospective cohort study
MMAS	163/652	USA	1987-1989	41-70	9 (7)	Prospective cohort study of the physiology of ageing
NSHDC E2	280/554	Sweden	1987-2000	40-62	4.5 (5)	Combination of intervention and population-based monitoring study of cardiovascular disease
NSHDC IGFs	281/575	Sweden	1987-2000	40-73	4 (4)	
NSHDC androg	407/396	Sweden	1987-2001	39-61	6 (5)	
PCPT	1,044/1,044	USA and Canada	1993-1998	55-83	7 (2)	Randomised, placebo-controlled trial of finasteride and prostate cancer
PHS	1,084/2,357	USA	1982-1983	39-84	9 (5)	Randomised trial of aspirin and β-carotene among physicians
PLCO	1,314/1,801	USA	1993-2000	54-75	2 (2)	Randomised controlled multicentre trial for early detection of cancer of the prostate, lung, colorectum and ovary
ProtecT feas	283/806	UK	1999-2001	50-70	0 (0)	Cross-sectional study within the first phase of population-based PSA testing and randomised, controlled trial of treatments of localised prostate cancer

Continued on next page

Table 3.1 – continued from previous page

Study	Cases/controls	Country	Year of blood collection	Age of blood collection	Median years from blood collection to diagnosis (IQR)	Study design
ProtecT main	2,651/3,233	UK	2001-2009	45-71	0 (0)	Cross-sectional study within a population-based PSA testing and randomised, controlled trial of treatments of localised prostate cancer
RBS	120/353	USA	1984-1986	47-86	7 (7)	Prospective cohort study of residents of Rancho Bernardo, Southern California
SU.VI.MAX-I	100/400	France	1994-1995	43-61	6 (4)	Population-based, double-blind, placebo-controlled, randomized trial of supplementation with antioxidant vitamins and minerals (vitamin C, α-tocopherol, β-carotene, selenium, and zinc)
SU.VI.MAX-II	184/368	France	1994-1995	42-61	11 (2)	

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; BUPA=British United Provident Association Study; CARET=Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; CHS=Cardiovascular Health Study; CLUE=Campaign Against Cancer and Stroke ("Give Us a Clue to Cancer") Study; EPIC=European Prospective Investigation into Cancer and Nutrition; ERSPC=European randomized screening prostate cancer; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Health Survey; HIMS=Health In Men Study; HPFS=Health Professionals Follow-Up Study; IQR=interquartile range; JACC=Japan Collaborative Cohort Study for Evaluation of Cancer; Janus NBSBWG=Nordic Biological Specimen Biobank Working Group; JHCS=Japan-Hawaii Cancer Study; JPHC=Japan Public Health Center-based Prospective Study; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=The Prostate, Lung, Colorectal and Ovarian; ProtecT=Prostate Testing for Cancer and Treatment; RBS=Rancho Bernardo Study; SU.VI MAX=SUpplémentation en VItamines et Minéraux AntioXydants.

Participant data collection

Participant data collection and blood sampling was carried out by each individual study. At study baseline men provided blood samples as well as information including anthropometric, sociodemographic and behavioural factors. Prostate cancer case status was followed-up, generally via record linkage (Chapter 6).

Study pooling and harmonisation

After individual studies were contacted to request collaboration with the EHNBPCCG, the data requested included: anthropometric (including height, weight, waist circumference and waist-to-hip ratio (WHR)), behavioural (smoking and alcohol), sociodemographic (racial/ethnic group, education status), and health (family history of prostate cancer).

Blood sample information was also requested including assay information and hormone measurements:

- Sex hormones: androstenedione, androstanediol glucuronide (A-diol-G), androsterone glucuronide, dehydroepiandrosterone sulfate (DHEAS), dehydroepiandrosterone (DHEA), testosterone, free testosterone, dihydrotestosterone (DHT), estradiol, free estradiol, estrone and sex hormone binding globulin (SHBG).

- IGFs: IGF-I, IGF-II, IGFBP-1 (insulin-like binding protein-1), IGFBP-2, and IGFBP-3.

- Other: cortisol, thyroid stimulating hormone and PSA at blood collection.

- Nutritional biomarkers.

Disease characterisation data were also requested including age at diagnosis, and tumour stage and grade. The data from each study were harmonised and stored in a central database.

Matching criteria

Matching criteria varied between studies, but always included age, which is strongly correlated with prostate cancer incidence, and a limited range of other factors where the estimation of effect was not of interest, such as date of recruitment, time of blood collection and study centre. The case/control ratio varied from 1:1 to 1:4. Some studies used density sample, meaning that an individual could appear more than once in a data file.

3.2.5 Ethical approval

The EHNBPCCG dataset contains secondary data, therefore further ethical approval for re-analysis was not necessary. However, each study individually obtained ethical approval and further details of participant consent and study design can be found in the original publications.

3.3 English national health records

3.3.1 Introduction

The National Health Service (NHS) routinely collects individual patient data regarding hospital admissions, diagnoses and other information, which is stored in various databases for use by government departments, British national bodies and regulators. Data can also be made available for research purposes upon application [264].

3.3.2 Hospital Episode Statistics Admitted Patient Care (HES APC)

Coverage

All admissions to NHS hospitals in England are recorded using HES APC, this also covers admissions to independent hospitals funded by the NHS [264, 265]. Between 98-99% of hospital activity in England is thought to be covered by the NHS [266]. A hospital admission includes any visit to a hospital which requires a hospital bed. This includes emergency and planned admission, day cases, births and operations [267], but excludes accident and emergency (A & E) attendance or outpatient bookings, which are held in separate databases. All HES databases are collated by NHS Digital.

Data collation

Following discharge from the care of a consultant (known as a finished consultant episode), a discharge summary for the patient including diagnoses and procedures will be completed. This summary is then forwarded to a clinical coding department, where coders will enter the information into a local patient information database. These coders are trained to use standardised procedures to translate the information into standardised diagnostic codes [268]. These local databases are pooled monthly by NHS Digital, who carry out data cleaning, remove duplicates and add geographic fields based on patient postcodes. NHS digital also attach pseudonymized patient identifiers to each admission (HESIDs) [269]. Each monthly dataset is collated and made available annually.

Patient information

This dataset also contains basic patient demographic information, including sex and postcode [270]. Individual-level measures of socioeconomic status are not available, but patient postcodes are used to estimate sociodemographic status using the Index of Multiple Deprivation (IMD). This index is a weighted area level aggregation of specific dimensions of deprivation including income, employment, health, education, barriers to housing and services, crime and living environment deprivation [271].

Diagnoses are coded using the International Classification of Disease 10^{th} revision (ICD-10) codes [268]. The number of diagnoses that it was possible to record per episode has increased over time from 7 before April 2002, 14 April 2002-March 2007, to 20 since April 2007. Each record contains one primary diagnosis, which accounts for the majority of the hospital admissions, while the other diagnoses are recorded as co-morbidities [267].

Record linkage

HES records are linked using the HESIDs so that individuals can be traced through multiple episodes over time [264, 265, 269]. HES APC is also routinely linked to the Register of Deaths in England and Wales (for deaths registered since 1998) [272] (see below). HESID missed match-rate is estimated to be $\sim$4% [273], which may be due to errors in the recording of NHS numbers [267].

Death registration

The Register of Deaths in England and Wales is recorded by the Office for National Statistics (ONS). Most deaths are certified by a medical practitioner, using the Medical Certificate of Cause of Death (MCCD) (Figure 3.1). Death registration records contain information both on primary cause of death and significant conditions contributing to the death but which are not related to the disease or condition causing it [274].

Patient consent

Prior to 11th October 2018, patients who did not wish their data to be shared beyond NHS Digital were able to opt-out at their GP practice [275]. This has since been replaced by a national data opt-out, which is declared using an online service [276]. The latest available estimate (March 2019) indicates 2.7% of registered patients have opted-out [277].

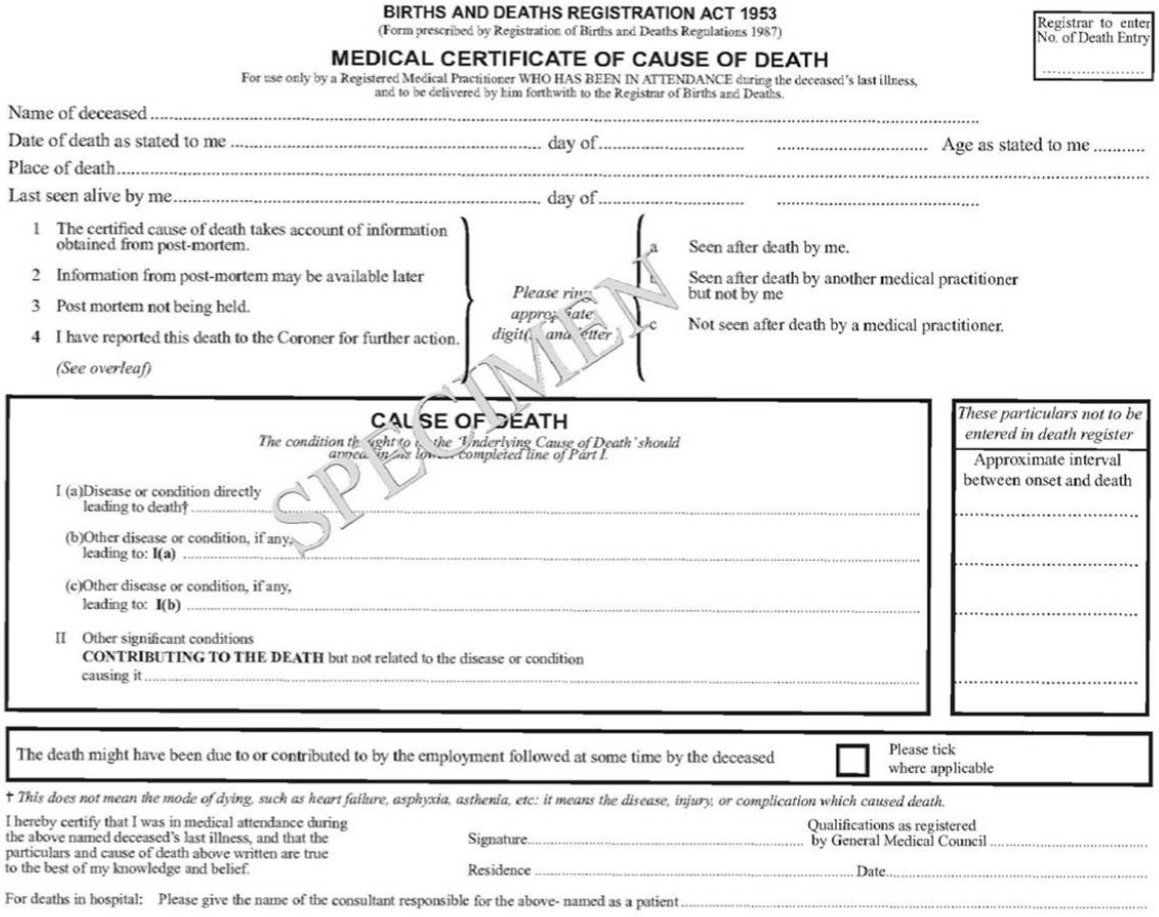
3.3.3 Ethical approval

The construction, maintenance, and analysis of the dataset for research were approved by the Central and South Bristol Research Ethics Committee (04/Q2006/176). The committee determined that individual patient consent was not required on the basis that this was a study without patient recruitment and used routine administrative data, in which the records contained no personally identifiable information or were pseudonymized through the encryption of personal identifiers.

Key contributors to the HES APC dataset and analysis

HES APC data is maintained and cleaned by NHS Digital who also routinely link the data to national mortality records, NHS Digital have granted permission for the Unit of Health Care Epidemiology to store and analyse the data.

Making these data available to the Cancer Epidemiology Unit would have expended substantial resources. Therefore, Raphael Goldacre, Unit of Health Care Epidemiology, conducted the statistical analysis of this dataset. I developed the analysis plan and created the tables and figures.



BIRTHS AND DEATHS REGISTRATION ACT 1953
(Form prescribed by Registration of Births and Deaths Regulations 1987)

Registrar to enter No. of Death Entry .....

MEDICAL CERTIFICATE OF CAUSE OF DEATH

For use only by a Registered Medical Practitioner WHO HAS BEEN IN ATTENDANCE during the deceased's last illness, and to be delivered by him forthwith to the Registrar of Births and Deaths.

Name of deceased .....

Date of death as stated to me ..... day of .... Age as stated to me .....

Place of death .....

Last seen alive by me ..... day of ....

1 The certified cause of death takes account of information obtained from post-mortem.
2 Information from post-mortem may be available later
3 Post mortem not being held.
4 I have reported this death to the Coroner for further action.
(See overleaf)

Please ring appropriate digit(s) and letter

a Seen after death by me.
b Seen after death by another medical practitioner but not by me
c Not seen after death by a medical practitioner.

SPECIMEN

CAUSE OF DEATH

The condition thought to be the 'Underlying Cause of Death' should be entered on the lowest completed line of Part I.

These particulars not to be entered in death register

Approximate interval between onset and death

I (a) Disease or condition directly leading to death† .....
(b) Other disease or condition, if any, leading to: **I(a)** .....
(c) Other disease or condition, if any, leading to: **I(b)** .....

II Other significant conditions **CONTRIBUTING TO THE DEATH** but not related to the disease or condition causing it .....

The death might have been due to or contributed to by the employment followed at some time by the deceased ☐ Please tick where applicable

† *This does not mean the mode of dying, such as heart failure, asphyxia, asthenia, etc: it means the disease, injury, or complication which caused death.*

I hereby certify that I was in medical attendance during the above named deceased's last illness, and that the particulars and cause of death above written are true to the best of my knowledge and belief.

Signature ..... Qualifications as registered by General Medical Council .....

Residence ..... Date .....

For deaths in hospital: Please give the name of the consultant responsible for the above- named as a patient .....

Figure 3.1: Medical Certificate of Cause of Death.
Source:[274].

3.4 UK Biobank

3.4.1 Introduction

The UK Biobank is a large prospective cohort study. The study was established with the aim of improving the prevention, diagnosis and treatment of a wide range of illnesses in middle-aged men and women [278].

Key contributors to the UK Biobank dataset and analysis

UK Biobank was founded by Professor Rory Collins, Clinical Trial Service Unit, University of Oxford. Participant data were collected and stored centrally. Upon application a subset of the data has been made available to collaborators within the Cancer Epidemiology Unit for which Associate Professor Ruth Travis is the Principle Investigator. I led and carried out the statistical analysis of the UK Biobank dataset for this thesis.

3.4.2 Design and methods

UK Biobank consists of approximately 500,000 individuals, recruited between 2006 and 2010. All participants were registered with the NHS. At baseline participants provided data on a range of sociodemographic, anthropometric, diet and lifestyle factors. Biological samples were also collected across the entire cohort and have been used to conduct assays and genotype information [279].

All participants are followed-up for health outcomes through record linkage. For subsets of the cohort, participants also provided data on diet, lifestyle and anthropometric factors at regular intervals.

Recruitment

Approximately 9.2 million individuals registered with the NHS aged between 40-69 years and located within 40 km of one of 22 recruitment centres located within England, Scotland and Wales, were invited to participate in UK Biobank via postal invitation [279]. 503,317 individuals consented to participate, resulting in a participation rate of 5.5%. The distribution of the assessment centres aimed to

provide socioeconomic and ethnic heterogeneity as well as a urban-rural mix, with the aim of being able to estimate reliable and generalisable associations between exposures and diseases [279].

Data collection

Sociodemographic, lifestyle and health history data

Data on participant lifestyle and other potentially health-related factors were collected at baseline via self-completed touch-screen questionnaires completed at the assessment centres. Socioeconomic status was assessed using factors such as housing tenure, car ownership, household income and structure, employment and income and education [279].

Health behaviour information was also collected, including smoking (current, previous, duration, frequency and cessation). Alcohol consumption was investigated using similar parameters and included the additional identification of beverage-specific drinking behaviour to improve the accuracy of estimations of alcohol consumption [279].

A computer-assisted personal interview was conducted to collect information which was not easily collected by touchscreen. Prior to these interviews, an aide was provided to participants to assist with the recollection of various types of information (such as medication, family history of disease, birth details). Questions at this interview were selected based on responses to the touch-screen questionnaire (e.g. if the participant reporting previous medical conditions, this prompted further disease-specific follow-up questions). The interview included a broad range of other questions relating to early-life (birth weight, breastfeeding, childhood size, place of birth), reproductive history, health history of first degree relatives, medical history, and further phenotype information [279].

Physiological measurements

Physiological measurements were collected at baseline across the whole cohort. Blood pressure was measured using the Omron HEM-7015IT digital blood pressure monitor.

Readings were taken twice with a one-minute rest in between measurements. Weight and bio-impedance were measured using the Tanita BC-418 MA body composition analyser, which can measure weight accurately to within 0.1 kg; participants were requested to remove shoes and heavy outer clothing prior to being weighed. Standing and sitting height were measured using a Seca 202 height measure. Hip and waist (at the level of the umbilicus) circumference were measured with a Wessex non-stretchable sprung tape measure [279].

Blood samples

40-50 ml of blood was taken during the baseline assessment visit using a vacutainer system [280]. Blood sampling was performed by either a phlebotomist or a nurse, in all participants except for a small proportion (0.3%) who declined, were deemed unable to, or where the attempt was abandoned for either technical or health reasons.

Blood was divided into Ethylenediaminetetraacetic acid (EDTA) vacutainers (for plasma samples and haematological assays), serum vacutainers (serum samples) and acid citrate dextrose (for functional assays). A urine sample as also collected and stored in six 1.4 ml aliquot tubes.

Samples were shipped to the central processing laboratory in temperature-controlled shipping boxes and maintained at 4°C. On arrival, the biological samples which were not used for the haematological assays (see below) were frozen. Following completion of recruitment, UK Biobank stored about 15 million 1.4 ml aliquot tubes in two separate locations. A "working" archive contains approximately 9 million samples at -80°C, while a "back-up" archive holds approximately 6 million samples preserved in liquid nitrogen at -196°C [280].

Haematological assays: Complete blood count and reticulocyte parameters were measured using a Beckman automated haematology analyser, typically within 24 hours of blood draw at the central laboratory [280]. A maximum of 31 haematological parameters were measured by the machine, including a mixture of measured, derived and calculated values [281] (see Chapter 8 for further details).

Blood biomarker panel: UK Biobank was due to measure a range of biomarkers (including testosterone and IGF-I) from some of these stored samples. However, some participant serum samples were inadvertently diluted during the creation of the aliquots from the serum vacutainer [282–284]. Correction for this error meant that the data release of this biomarker panel was delayed until March 2019 and therefore, these measurements were not available at the time of data analyses for this thesis.

Genomic data: Genotyping was carried out on participants using blood collected at baseline. Initially, a subset of 49,950 participants were genotyped at 807,411 markers using the Applied Biosystems UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) Axiom Array by Affymetrix (now part of Thermo Fisher Scientific) [285]. Later, 483,427 participants were genotyped using the Applied Biosystems UK Biobank Axiom Array, with 825,927 markers (this shares 95% of marker content with the UK BiLEVE axiom array [286]). Genotyping was carried out by Affymetrix Research Services Laboratory in batches of approximately 4,700 samples [285].

Dietary questionnaire data

At baseline all participants self-completed a short set of food frequency questions, containing 29 questions about diet and 18 about alcohol. This touch-screen questionnaire collected data on the frequency of consumption over the past year of food groups, avoidance of foods, age last ate meat and changes in diet over the past five years [287].

Diet was assessed in greater detail in a subset of participants using a web-based 24-hour dietary assessment tool. Participants who were recruited between April 2009 and September 2010 completed the Oxford WebQ at the assessment centre [288] (n=70,046 participants) [289]. This questionnaire asked questions about the consumption of up to 206 foods and 32 types of drinks during the previous 24 hours.

Participants were asked a series of yes/no questions about whether they ate specific foods. A positive response led to follow-up questions about quantity/portion sizes. Mean daily intakes of nutrients were calculated by multiplying the frequency of consumption of each food or drink by the quantity consumed and the nutrient composition of each item, taken from the UK food composition database McCance and Widdowson's The Composition of Foods and its supplements.

After the recruitment period closed, all participants who provided an email address at recruitment were emailed every 3–4 months a total of four times between February 2011 and June 2012, inviting them to complete the Oxford WebQ online using their own computers [288]. 210,126 participants completed have at least one 24-h dietary assessment [289].

Data storage and anonymity

UK Biobank does not provide identifiable data such as NHS number, name or postal address. Data are provided to researchers on an application-specific basis, via a reverse-anonymised form encoded using an irreversible algorithm. UK Biobank retains an internal database which enables only UK Biobank to reverse anonymise the data [287].

3.4.3 Ethical approval

All participants provided written informed consent. Ethical approval for UK Biobank was obtained by the North West Multi-Centre Research Ethics Committee for local NHS Primary Care Trusts, the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland.

Research in this thesis was conducted using the UK Biobank resource application number 3282.

3.4.4 Follow-up

Further assessments

Following the baseline assessment subsets of approximately 20,000 participants are scheduled to be contacted every 2-3 years during follow-up for repeat assessments to provide information on participant lifestyle, health and biomarkers over time and also allows for the correction of regression dilution bias [290, 291]. Some further exposures are also evaluated in greater depth including physical activity and diet in subsets of participants [287].

Disease incidence and mortality

UK Biobank can obtain access to previous and future medical record data for all participants via their NHS numbers in England and Wales, and Community Health number in Scotland. This enables researchers to track participant health outcomes via access to databases including death and cancer registries, general practice and hospital activity records [279].

4

Cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating sex hormone concentrations in men

Contents

4.1 Introduction 72
4.1.1 Aims 73
4.2 Subjects and methods 73
4.2.1 Data collection 73
4.2.2 Data processing 80
4.2.3 Statistical analysis 82
4.2.4 Multiple testing 84
4.2.5 Statistical software 84
4.3 Results 84
4.3.1 Correlations between hormones 84
4.3.2 Age at blood collection 86
4.3.3 Adiposity 89
4.3.4 Height 93
4.3.5 Smoking 93
4.3.6 Alcohol 93
4.3.7 Ethnic/racial group 97
4.3.8 Time of blood collection 97
4.3.9 Sociodemographic factors and health 97
4.3.10 Heterogeneity by study and assay 103
4.3.11 Sensitivity analyses 103
4.4 Discussion 104

4.4.1 Correlations between hormones 104
4.4.2 Age 104
4.4.3 Adiposity 105
4.4.4 Height 106
4.4.5 Smoking 106
4.4.6 Alcohol 106
4.4.7 Racial/ethnic group 107
4.4.8 Heterogeneity between studies 107
4.4.9 Calculation of free testosterone and estradiol 107
4.4.10 Strengths and limitations 108
4.5 Conclusion 109

4.1 Introduction

Sex hormones are fundamental to many aspects of physiology. In the prostate, androgens modulate cell proliferation, differentiation and apoptosis and have been theorised to be involved in prostate cancer development. Although prospective epidemiological studies have not found evidence of an association between circulating sex hormone concentrations and prostate cancer risk [13], clinical and experimental evidence indicates that there may still be an association between circulating sex hormones and prostate cancer (Chapter 2; also see analyses in Chapters 6 and 7). In addition to their putative role in prostate cancer, sex hormones have been implicated in the development of other chronic diseases, including hypertension, diabetes, cardiovascular and renal disease [292–296].

Circulating sex hormone concentrations may be partially modified by a range of anthropometric, sociodemographic and behavioural factors [297–299]. Although some of these associations have been previously investigated [297–301], fewer studies have examined these factors in relation to androstenedione, androstanediol glucuronide (A-diol-G) and estrone, and individual studies are generally underpowered to investigate associations at extremes of the distributions.

4.1.1 Aims

Using the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG) dataset, this chapter aims to examine the anthropometric, sociodemographic, and behavioural correlates of sex hormones and sex hormone binding globulin (SHBG). Understanding these associations will lead to a better understanding of plausible biological mechanisms of how these factors may affect disease risk.

Using unified methods, this large individual-level participant dataset is well placed to describe associations of sex hormones at extremes of their distribution and for most analyses, with substantial statistical power to identify moderate magnitude but novel associations.

4.2 Subjects and methods

4.2.1 Data collection

The EHNBPCCG search strategy and pooling is described in greater detail in Chapter 3. In brief, principal investigators were invited to contribute data to the EHNBPCCG if they had published or unpublished data on concentrations of endogenous hormones and/or nutritional biomarkers from blood samples collected from men prior to diagnosis of prostate cancer and male controls.

Collaborators provided data on a range of baseline sex hormones, including concentrations of androstenedione, A-diol-G, dehydroepiandrosterone sulphate (DHEAS), testosterone, free testosterone, dihydrotestosterone (DHT), estradiol, free estradiol, estrone and SHBG, together with assay details (Tables 4.1-4.3). Collaborators also provided data on a range of exposures: anthropometric (including height, weight, waist circumference and waist-to-hip ratio (WHR)), behavioural (smoking and alcohol), sociodemographic factors (racial/ethnic group, education status), and health (family history of prostate cancer), generally collected at the same time as blood collection. The data from each study were harmonised and stored in a central database.

Table 4.1: Assay methods and geometric mean sex hormone and SHBG concentrations.

	Androstenedione (nmol/L)			A-diol-G (nmol/L)		
	Method	CV %	Geometric mean (95% CI)	Method	CV %	Geometric mean (95% CI)
ATBC	E RIA	7†	4.7 (4.4-4.9)	E RIA	12.3†	5.8 (5.5-6.2)
BLSA	-	-	-	-	-	-
CARET	E RIA	5–9‡	2.7 (2.6-2.8)	NE RIA	1.5-4.8‡	12.5 (12.0-13.1)
CHDS	-	-	-	-	-	-
EPIC phase I	NE RIA	3.5-11.1†	4.6 (4.5-4.7)	NE RIA	4.1-9.9†	12.9 (12.5-13.3)
EPIC Norfolk	-	-	-	N/S		12.0 (11.5-12.5)
FMC	NE RIA	9.5-11.7[a]	5.9 (5.6-6.1)	-	-	-
HHS NBSBWG	-	-	-	-	-	-
HPFS I	-	-	-	NE RIA	6.7†	10.9 (10.5-11.2)
JACC	-	-	-	-	-	-
Janus	-	-	-	-	-	-
JHCS	E RIA	5-15‡	5.0 (4.7-5.2)	E RIA	5-15‡	9.3 (8.7-10.0)
JPHC	-	-	-	-	-	-
KPMCP	-	-	-	-	-	-
MCCS	NE RIA	3.3[b]	3.5 (3.5-3.6)	NE RIA	4.3[b]	13.6 (13.2-14.0)
MEC	-	-	-	NE RIA	3.4†	9.5 (9.2-9.8)
MMAS	E RIA	8.6-10‡	2.9 (2.8-3.0)	-	-	-
NSHDC	-	-	-	N/S		11.7 (11.3-12.2)
PCPT	E RIA	6.9-8.6[a]	2.0 (1.9-2.1)	NE RIA	2.7-14.0‡	12.0 (11.7-12.3)
PHS	-	-	-	NE RIA	7.6†	13.8 (13.2-14.5)
PLCO	NE RIA	14†	4.1 (4.0-4.2)	NE RIA	11†	14.2 (13.8-14.6)
ProtecT feas	-	-	-	-	-	-
RBS	-	-	-	-	-	-

†Intra-assay.

[a]Inter-assay.

‡Intra-and inter-assay range.

[b]Not specified.

Abbreviations: A-diol-G=androstanediol glucuronide; ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; CI=confidence interval; CIA=competitive immunoassay; CV=coefficient of variation; E RIA=extraction radioimmunoassay; ECIA=electrochemiluminescence immunoassay; EIA=enzyme immunoassay; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS NBSBWG=Helsinki Heart Study Nordic Biological Specimen Biobank Working Group; HPFS=Health Professionals Follow-Up Study; IA=immunoassay; IMF=immunofluorimetry; IRMA=immunoradiometric assay; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; JHCS=Japan-Hawaii Cancer Study; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NE RIA=non-extraction radioimmunoassay; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; ProtecT feas=Prostate Testing for Cancer and Treatment feasibility study; RBS=Rancho Bernardo Study.

Table 4.2: Assay methods and geometric mean sex hormone and SHBG concentrations (continued).

	DHEAS (nmol/L)			DHT (nmol/L)			Testosterone (nmol/L)		
	Method	CV (%)	Geometric mean (95% CI)	Method	CV %	Geometric mean (95% CI)	Method	CV %	Geometric mean (95% CI)
ATBC	E RIA	11.3†	3092 (2886-3313)	E RIA	8.9†	1.82 (1.72-1.93)	E RIA	5.5†	20.3 (19.5-21.2)
BLSA	NE RIA	2-6.6‡	3337 (3058-3641)	-	-	-	NE RIA	3.3-6.4‡	15.1 (14.4-15.9)
CARET	NE RIA	0.8-1.9‡	1994 (1892-2101)	-	-	-	E RIA	1-12‡	13.8 (13.4-14.2)
CHDS	-	-	-	-	-	-	E RIA	9-11[a]	22.1 (21.6-22.6)
EPIC phase I	-	-	-	-	-	-	NE RIA	10.8-14.8†	15.8 (15.5-16.2)
EPIC Norfolk	N/S		2417 (2317-2522)	-	-	-	N/S		15.2 (14.8-15.6)
FMC	-	-	-	-	-	-	NE RIA	4.5-7.2[a]	23.1 (22.3-24.0)
HHS NBSBWG	-	-	-				IMF	5.5-13‡	19.8 (19.0-20.6)
HPFS I	-	-	-	NE RIA	9.7†	1.22 (1.18-1.27)	ECIA	4.9†	15.5 (15.1-15.8)
JACC	-	-	-	-	-	-	NE RIA	5-12[b]	15.8 (14.8-16.8)
Janus	-	-	-	E RIA	5-15‡	1.56 (1.46-1.66)	E RIA	5-15‡	18.7 (17.8-19.6)
JHCS	-	-	-	E RIA	5-15‡	1.81 (1.70-1.92)	E RIA	5-15‡	17.3 (16.6-18.1)
JPHC	-	-	-	-	-	-	ECIA	1-3‡	15.6 (15.1-16.1)
KPMCP	NE RIA	N/S	2533 (2340-2742)	-	-	-	-	-	-
MCCS	CIA	12.4[b]	2703 (2618-2791)	-	-	-	ECIA	1.6[b]	15.3 (15.1-15.6)
MEC	-	-	-	E RIA	3.8†	1.92 (1.87-1.98)	E RIA	3.5†	18.3 (17.9-18.6)
MMAS	NE RIA	4.1-8.9‡	5489 (5248-5742)	NE RIA	10.9-12.2‡	0.75 (0.73-0.78)	E RIA	4.6-7.2‡	16.4 (16.0-16.9)
NSHDC	-	-	-	-	-	-	N/S		20.0 (19.4-20.5)
PCPT	-	-	-	-	-	-	ECIA	7.6-11.9‡	12.5 (12.3-12.7)
PHS	-	-	-	E RIA	5.3†	1.17 (1.12-1.22)	NE RIA	8.7†	16.0 (15.5-16.5)
PLCO	-	-	-	-	-	-	NE RIA	14†	16.0 (15.7-16.3)
ProtecT feas	-	-	-	-	-	-	EIA	N/S	13.9 (13.6-14.3)
RBS	NE RIA	3.1-7.3‡	2088 (1959-2226)	E RIA	7.5‡	1.39 (1.32-1.46)	E RIA	4.1-10‡	10.2 (9.9-10.6)

†Intra-assay.

[a]Inter-assay.

‡Intra-and inter-assay range.

[b]Not specified.

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; CI=confidence interval; CIA=competitive immunoassay; CV=coefficient of variation; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; E RIA=extraction radioimmunoassay; ECIA=electrochemiluminescence immunoassay; EIA=enzyme immunoassay; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS NBSBWG=Helsinki Heart Study Nordic Biological Specimen Biobank Working Group; HPFS=Health Professionals Follow-Up Study; IA=immunoassay; IMF=immunofluorimetry; IRMA=immunoradiometric assay; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; JHCS=Japan-Hawaii Cancer Study; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NE RIA=non-extraction radioimmunoassay; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; ProtecT feas=Prostate Testing for Cancer and Treatment feasibility study; RBS=Rancho Bernardo Study.

Table 4.3: Assay methods and geometric mean sex hormone and SHBG concentrations (continued).

	Estradiol (pmol/L)			Estrone (pmol/L)			SHBG (nmol/L)		
	Method	CV %	Geometric mean (95% CI)	Method	CV (%)	Geometric mean (95% CI)	Method	CV %	Geometric mean (95% CI)
ATBC	E RIA	13.7†	102 (99-106)	E RIA	14.1†	153 (148-159)	IRMA	4.2†	84.4 (80.4-88.5)
BLSA	-	-	-	-	-	-	NE RIA	1.8-22‡	78.5 (74.0-83.3)
CARET	E RIA	5-13‡	191 (186-196)	-	-	-	IRMA	2-7‡	27.3 (26.3-28.3)
CHDS	E RIA	8-10[a]	161 (158-165)	-	-	-	IA	4-6[a]	32.9 (32.0-33.8)
EPIC phase I	-	-	-	-	-	-	IRMA	7.7-12.2†	42.0 (41.0-43.1)
EPIC Norfolk	-	-	-	-	-	-	N/S		41.9 (40.6-43.3)
FMC	-	-	-	-	-	-	IMF	6.6-8.7[a]	50.8 (48.8-53.0)
HHS NBSBWG	-	-	-	-	-	-	IMF	1.3-10.1‡	50.3 (48.0-52.6)
HPFS I	NE RIA	5.2†	111 (109-113)	-	-	-	IRMA	10.7†	56.9 (55.5-58.3)
JACC	-	-	-	-	-	-	IRMA	5.6-6.9[b]	42.8 (39.6-46.1)
Janus	-	-	-	-	-		-	-	-
JHCS	E RIA	9-15‡	65 (61-68)	E RIA	9-15‡	106 (101-111)	Precipitation	5-15‡	37.4 (35.0-39.8)
JPHC				-	-	-	IRMA	2-8‡	47.0 (45.3-48.8)
KPMCP				-	-	-	-	-	-
MCCS	ECIA	11.1[b]	106 (105-108)	-	-	-	IA	6[b]	36.8 (36.0-37.7)
MEC	-	-	-	-	-	-	ECIA	3†	35.7 (34.9-36.6)
MMAS	NE RIA	3.6-7.1‡	138 (135-141)	-	-	-	Filtration assay	8-10.9‡	30.4 (29.4-31.4)
NSHDC	N/S	-	76 (74-78)	-	-	-	IRMA	N/S	41.6 (40.2-42.9)
PCPT	E RIA	10-14.9‡	119 (117-121)	E RIA	9-15.2‡	159 (157-162)	ECIA	5.2-12.2‡	36.6 (35.9-37.3)
PHS	-	-	-	E RIA	10†	120 (117-124)	IRMA	8.9†	20.3 (19.5-21.0)
PLCO	-	-	-	-	-	-	IRMA	18†	43.5 (42.5-44.5)
ProtecT feas	-	-	-	-	-	-	EIA	-	37.1 (36.0-38.3)
RBS	E RIA	8-12‡	72 (70-75)	-	-	-	-	-	-

†Intra-assay.

[a]Inter-assay.

‡Intra-and inter-assay range.

[b]Not specified.

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; CI=confidence interval; CIA=competitive immunoassay; CV=coefficient of variation; E RIA=extraction radioimmunoassay; ECIA=electrochemiluminescence immunoassay; EIA=enzyme immunoassay; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS NBSBWG=Helsinki Heart Study Nordic Biological Specimen Biobank Working Group; HPFS=Health Professionals Follow-Up Study; IA=immunoassay; IMF=immunofluorimetry; IRMA=immunoradiometric assay; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; JHCS=Japan-Hawaii Cancer Study; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NE RIA=non-extraction radioimmunoassay; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; ProtecT feas=Prostate Testing for Cancer and Treatment feasibility study; RBS=Rancho Bernardo Study; SHBG=sex hormone binding globulin.

Inclusion criteria

Participants were considered eligible for this analysis if they had information on at least one of the circulating sex hormones listed above or SHBG, had not been diagnosed with prostate cancer by the time of the individual study closure, were 25 years or over at blood collection (to ensure these men are post-puberty), had complete data on age at blood collection, height, and weight (to enable the calculation of BMI), and were not known to be receiving androgen therapy at blood collection.

Overall, these inclusion criteria resulted in 12,330 participants (out of a total of 14,092; Figure 4.1) from 25 studies being included in these analyses.

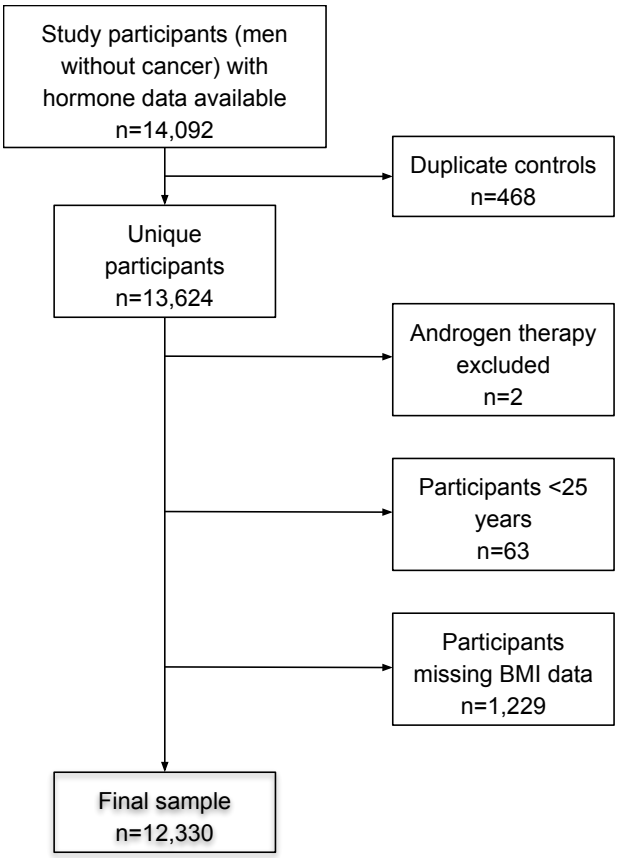


Figure 4.1: Participant selection chart.
Abbreviations:BMI=body mass index.

Included studies

The following studies were included in this analysis: Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC) [229], Baltimore Longitudinal Study of Aging (BLSA) [230, 231], Carotene and Retinol Efficacy Trial (CARET) [233], Child Health and Development Studies (CHDS) [234], European Prospective Investigation into Cancer and Nutrition (EPIC) [237], EPIC Norfolk [238], Finnish Mobile Clinic (FMC) [240], Helsinki Heart Study (HHS) [243, 244], Health Professionals Follow-up Study (HPFS) [245], Japan Collaborative Cohort Study (JACC) [246], Japan Public Health Center-based prospective study (JPHC) [249], Janus serum bank [244, 250], Melbourne Collaborative Cohort Study (MCCS) [252], Multiethnic Cohort (MEC) [254], Massachusetts Male Aging Study (MMAS) [255], Northern Sweden Health and Disease Cohort (NSHDC) [244], Prostate Cancer Prevention Trial (PCPT) [10], Physicians' Health Study (PHS) [256], the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) [257], Prostate Testing for Cancer and Treatment (ProtecT) feasibility study [258–260], and the Rancho Bernado Study (RBS) [261] by dataset closure on 1st January, 2017.

Participant characteristics by study are summarised in Table 4.4. All studies were either of a prospective cohort design or prospective observational studies within a randomised trial. Further details of participant consent and study design can be found in the original publications and in Chapter 3.

Table 4.4: Participant characteristics by study.

Study, country	N (% of total)	Mean age, years (SD)	Age range, yrs	Year of birth	Mean height, cm (SD)	Mean BMI, kg/m^2 (SD)	% Current drinkers (median alcohol consumption, g ethanol/day)	% Current smokers (median cigarettes/ day)	% Non-Hispanic white ethnic group	% Married/ cohabiting at blood collection	% University degree	% Family history of prostate cancer
ATBC, Finland	231 (1.9%)	60.5 (5.2)	50-69	1915-1936	173.8 (6.8)	25.9 (3.7)	85.4 (8.5)	100 (20)	-	82.3	5.6	2.5
BLSA, USA	112 (0.9%)	56.4 (12.8)	30-84	1895-1948	176.8 (6.0)	25.7 (3.0)	-	21.4	95.5	91.8	57.1	-
CARET, USA	300 (2.5%)	62.8 (6.0)	47-77	1913-1945	174.9 (7.6)	27.7 (4.4)	68.2 (5)	50.0 (20)	93.7	85.9	24.2	3.3
CHDS, USA	396 (2.4%)	34.9 (3.2)	25-50	1912-1941	178.7 (6.8)	24.7 (2.7)	77.5 (6)	50.6 (20)	63.1	99.7	34.5	-
EPIC phase I, Europe	638 (5.2%)	60.9 (6.2)	43-76	1918-1952	172.6 (7.0)	27.0 (3.6)	87.1 (13)	27.8 (15)	100	89.1	22.9	-
EPIC Norfolk, Europe	707 (5.7%)	66.7 (6.5)	45-78	1918-1950	172.9 (6.4)	26.6 (3.2)	73.8 (8)	9.3	99.9	88.3	12.5	1.7
FMC, Finland	298 (2.4%)	57.9 (9.9)	33-80	1890-1937	169.6 (6.8)	26.1 (3.5)	-	34.2 (15)	100	85.2	-	-
HHS NBSBWG, Finland	291 (2.4%)	51.0 (3.8)	40-56	1924-1940	174.2 (6.2)	26.4 (3.1)	-	-	-	-	-	-
HPFS I, USA	682 (5.5%)	65.1 (7.4)	46-80	1913-1946	178.1 (6.4)	26.0 (3.5)	73.5 (6)	3.9 (20)	99.4	93	100	10.3
JACC, Japan	97 (0.8%)	67.8 (5.8)	58-83	1904-1934	159.5 (7.1)	22.4 (2.7)	52.6 (2)	37.2 (20)	0	93.8	2.7	-
Janus, Norway	1,153 (9.4%)	46.4 (4.2)	33-63	1910-1947	176.2 (6.8)	25.1 (3.2)	-	62.5 (12)	100	-	-	-
JHCS 1988, Japanese in Hawaii	98 (0.8%)	63.4 (5.6)	52-74	1900-1919	162.2 (5.4)	23.1 (2.8)	74.5 (5.5)	36.7 (20)	0	90.8	8.2	-
JHCS 1996, Japanese in Hawaii	139 (1.1%)	61.9 (5.7)	52-74	1900-1919	162.8 (6.4)	23.2 (2.8)	78.4 (5)	28.1 (20)	0	94.2	10.1	-
JPHC, Japan	399 (3.2%)	59.1 (6.6)	41-71	1923-1949	162.0 (6.4)	23.3 (2.6)	71.2 (23)	40.6 (20)	0	91	-	-
KPMCP, USA	212 (1.7%)	71.8 (4.5)	60-85	1882-1906	169.9 (6.7)	25.8 (3.1)	69.7 (10)	17.9 (30)	98.6	82.7	5.4	
MCCS, Australia	1,058 (8.6%)	58.3 (7.2)	40-72	1918-1954	172.2 (7.3)	27.2 (3.7)	78.3 (13)	13.0 (20)	100	81	22.2	-
MEC, USA	922 (7.5%)	68.7 (7.2)	48-84	1918-1948	174.2 (7.7)	27.1 (4.2)	90.9 (9)	12.6 (15)	13.2	78.1	32.2	9
MMAS, USA	651 (5.3%)	57.6 (7.0)	41-70	1916-1946	175.8 (7.2)	27.5 (4.5)	-	23.5	98.8	77.1	37.8	-
NSHDC E2, Sweden	465 (3.8%)	57.6 (4.0)	40-61	1927-1948	175.8 (5.9)	26.6 (3.5)	-	21.3	100	-	-	-
NSHDC androg, Sweden	382 (3.1%)	57.2 (4.5)	39-61	1927-1958	175.8 (6.2)	26.5 (3.4)	100.0 (4)	17.4	100	85.6	12.4	
PCPT, USA and Canada	1,025 (8.3%)	63.3 (5.5)	55-83	1911-1940	177.4 (7.0)	27.7 (4.0)	68.9 (3)	7.6 (20)	83.9	87.7	37.1	20.7
PHS, USA	376 (3.0%)	61.3 (7.5)	41-77	1905-1941	177.9 (7.0)	24.6 (2.5)	84.4 (5)	8.5 (20)	96.5	-	100	-
PLCO, USA	858 (7.0%)	64.8 (4.8)	54-75	1919-1944	177.9 (6.5)	27.4 (3.9)	70.0 (4)	9.1 (20)	100	86.9	42.2	6.1
ProtecT feas, UK	539 (4.4%)	61.5 (5.1)	50-70	1930-1949	175.2 (6.5)	26.6 (3.5)	84.9 (18)	10.5	-	-	-	4.1
RBS, USA	301 (2.4%)	69.4 (9.0)	47-85	1898-1938	174.4 (6.8)	26.0 (3.6)	68.8 (10)	7.3 (20)	100	90.4	22.6	-
Overall†	12,330	59.9 (9.9)	25-85	1882-1958	174.4 (7.8)	26.4 (3.8)	77.1 (8)	23.7 (20)	89.4	86.5	35.8	8.2
% Missing data[a]	-	-	-	-	-	-	31.1	12.2	9	26.3	27.3	55.5

†Excludes missing data.

[a]Excludes studies where data were not collected.

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS NBSBWG=Helsinki Heart Study Nordic Biological Specimen Biobank Working Group; HPFS=Health Professionals Follow-Up Study; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; JHCS=Japan-Hawaii Cancer Study; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; ProtecT feas=Prostate Testing for Cancer and Treatment feasibility study; RBS=Rancho Bernardo Study.

4.2.2 Data processing

Sex hormone and SHBG concentrations were log-transformed to approximate normal distributions. The analyses examined associations of these hormones and SHBG with age (25-49 (mean age=43.5), 50-54, 55-59, 60-64, 65-69, 70-74, 75+ years), body mass index (BMI (<20.0, 20.0-22.4, 22.5-24.9, 25.0-27.4, 27.5-29.9, 30.0-32.4, 32.5-34.9, 35.0-37.4, 37.5+ kg/m^2)), waist circumference (<90.0, 90.0-94.9, 95.0-99.9, 100.0-104.9, 105.0+ cm) and WHR (<0.900, 0.900-0.932, 0.933-0.966, 0.967-0.999, 1.000+), height (<160.0, 160.0-164.9, 165.0-169.9, 170.0-174.9, 175.0-179.9, 180.0-184.9, 185.0-189.9, 190.0+ cm), smoking status (never, former, current: <15, 15-29, 30+ cigarettes per day), alcohol consumption (none, 1-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70+ g ethanol per day), ethnic/racial group (non-Hispanic white, African American/Caribbean, Hispanic/Latino, East Asian, and other), time of day of blood collection (before 09:00, 09:00-11:59, 12:00-14:59, 15:00 onwards), educational status (no secondary/high school education, secondary school, college and university), marital status (currently married/cohabiting, or not currently married/cohabiting), and family history of prostate cancer (no, yes: defined as a father and/or brother diagnosed with prostate cancer). Categories of the exposure variables investigated were defined *a priori* based on sample size and the data distribution.

Free testosterone and free estradiol estimations

Free testosterone and free estradiol concentrations were estimated using equations based on the law of mass action [163, 302]. This law states that the rate of any chemical reaction is proportional to the product of the activities or concentrations of the reactants [303]. Therefore, if the concentrations and association constants under equilibrium conditions of the ligands and binding proteins are known then it will be possible to predict free hormone concentrations. This calculation assumes that all binding sites on each protein are equivalent and independent with no competition between hormones for the binding to albumin:

Free testosterone (pmol/L):

$$[FT] = 1.0\times10^{12}\times(\sqrt{((N/k_s - k_s + [S])^2/(4N^2) + k_s/(Nk_s))} - (N/k_s - [T] + [S])/(2N))$$

Where:

N = 1 + [A]k_A

[T] = total testosterone concentration (mol/L)

[S] = SHBG concentration (mol/L)

[A] = albumin concentration (assuming 6.2 x 10^{-4} mol/L)

k_s = association constant of testosterone to SHBG (assuming 1.0 x 10^9)

k_A= association constant of testosterone to albumin (assuming 3.6 x 10^4)

Free estradiol (pmol/L):

[FE] = 1.0 x 10^{12} $\times((-b + \sqrt{((b^2 - 4ac))})/(2a))$

Where:

a = (k_s x k_A [A]) + k_s

b = (([S]k_s) + ([A]k_A) - ([E]k_s)) + 1

c = -[E]

[E] = total estradiol concentration (mol/L)

[S] = SHBG concentration (mol/L)

[A] = albumin concentration (assuming 6.2 x 10^{-6} mol/L)

k_s = association constant of estradiol to SHBG (assuming 0.6 x 10^9)

k_A= association constant of estradiol to albumin (assuming 0.61 x 10^4)

4.2.3 Statistical analysis

Partial correlations between the sex hormones and SHBG were calculated using study-specific standardised values: $(x_{jk}-m_j)/s_j$, where m_j and s_j denote the mean and standard deviation of the log-transformed hormone concentrations in study j and x_{jk} is an observation from that study. These standardised values were adjusted for age at blood collection and BMI (included as categorical variables, described above).

Geometric mean concentrations of hormones/binding proteins were calculated using predicted values from analysis of variance models scaled to the overall geometric mean concentration and adjusted for study, age at blood collection, height, and BMI (with the exception of when the associations of age, height and BMI were analysed with hormone concentrations, when these variables were not included as adjustment covariates). Adjustment categories were chosen *a priori* on the basis of data availability. Adjusted geometric mean concentrations for waist circumference and WHR were also repeated with and without adjustment for BMI. When BMI and waist circumference are adjusted in the same model, BMI may reflect lean body mass more than overall fat mass [304]. Therefore, when waist is adjusted for BMI, this further adjustment may demonstrate the potential association of central and/or visceral adiposity with hormone concentrations [304].

Analyses of smoking and alcohol consumption were mutually adjusted for because these two lifestyle factors are often correlated [305] and studies which collected smoking information generally also collected alcohol consumption data.

To investigate the relationships of hormone concentrations with ethnicity/race, studies were limited to the three that had sufficient representation from men across multiple ethnic/racial groups (CHDS, MEC and PCPT).

Where appropriate, a test for trend was calculated using the analysis of variance test, with the categorical variables entered as linear values scored consecutively as 1, 2, 3 etc. Owing to the highly skewed distribution of alcohol consumption, the test for trend was calculated based on median values within each category. To test for trend by smoking status, never and former smokers were combined and coded as 0; light, medium and heavy smokers were coded as 1, 2 and 3, respectively, as current

smoking status may be more likely than previous smoking history to determine circulating sex hormone and SHBG concentrations. In a secondary analysis this test for trend by smoking status was calculated for current smokers only.

Most variables had a small number of missing values (Table 4.4). Time of day at blood collection was missing for 44.8% of men and was therefore not included as a covariate in the analyses. To enable adjustment for study, each study had to contain observations in a minimum of two categories for each primary exposure to be included in the respective exposure analysis.

Assay methods for each hormone varied between studies (Tables 4.1-4.3); some studies used "extraction" assays to measure circulating sex hormones (immunoassays preceded by extraction into an organic solvent and celite column chromatography), which generally provide higher sensitivity and specificity than non-extraction or "direct" assays. Heterogeneity between studies and by assay method was tested using a study-by-factor and assay-by-factor interaction term (fitted separately) in the analysis of variance and assessed using the F test. For the analysis of BMI, heterogeneity in the associations between the studies which used measured and self-reported BMI was assessed using the likelihood ratio test.

Sensitivity analysis

Separate analyses were performed stratified by: age (<55, 55-69, 70+ years); BMI (<20.0, 20.0-29.9, 30.0+ kg/m^2); and time of day (morning and afternoon), in comparison to a restricted dataset to examine the possibility of effect modification by these factors. The analyses were also repeated after restricting the dataset to: i) non-Hispanic white men only, and ii) men with sex hormone and SHBG concentrations that were within the range of [lower quartile – 3*interquartile range, upper quartile + 3*interquartile range] of each respective study in order to examine the effect of outliers (n=63). The primary analysis was also repeated after further adjustment for height, smoking and alcohol and using alternative calculations to estimate free testosterone and estradiol [306] (See appendix A for calculations).

4.2.4 Multiple testing

This analysis contains a large number of statistical tests, which increases the probability of type I errors. Because the analyses were exploratory, the analyses did not adjust further for multiple testing to avoid rejecting potentially important findings [307]. However, the associations observed, particularly those with weaker p-values, have been reported in the context of evidence from other studies.

4.2.5 Statistical software

All statistical tests were two-sided. Data analyses used Stata Statistical Software release 14.1 (Stata Corp., College Station, TX, USA), and figures were created in R version 3.2.3.

4.3 Results

25 studies contributed to the analysis, including 12,330 participants (Table 4.4). Participants were born between 1882-1958 and age at blood collection ranged from 25-85 years (mean=59.9 years; standard deviation=9.9 years). The men were predominantly non-Hispanic white (89.4%).

4.3.1 Correlations between hormones

After taking into account age, BMI and study, most sex hormones were positively correlated with each other (Figure 4.2). Testosterone was most strongly correlated with free testosterone (r=0.77), DHT (r=0.54) and SHBG (r=0.52). DHEAS was correlated with androstenedione (r=0.40) and A-diol-G (r=0.21). Estradiol was strongly correlated with free estradiol (r=0.88) and estrone (r=0.54), and weakly correlated with SHBG (r=0.10).

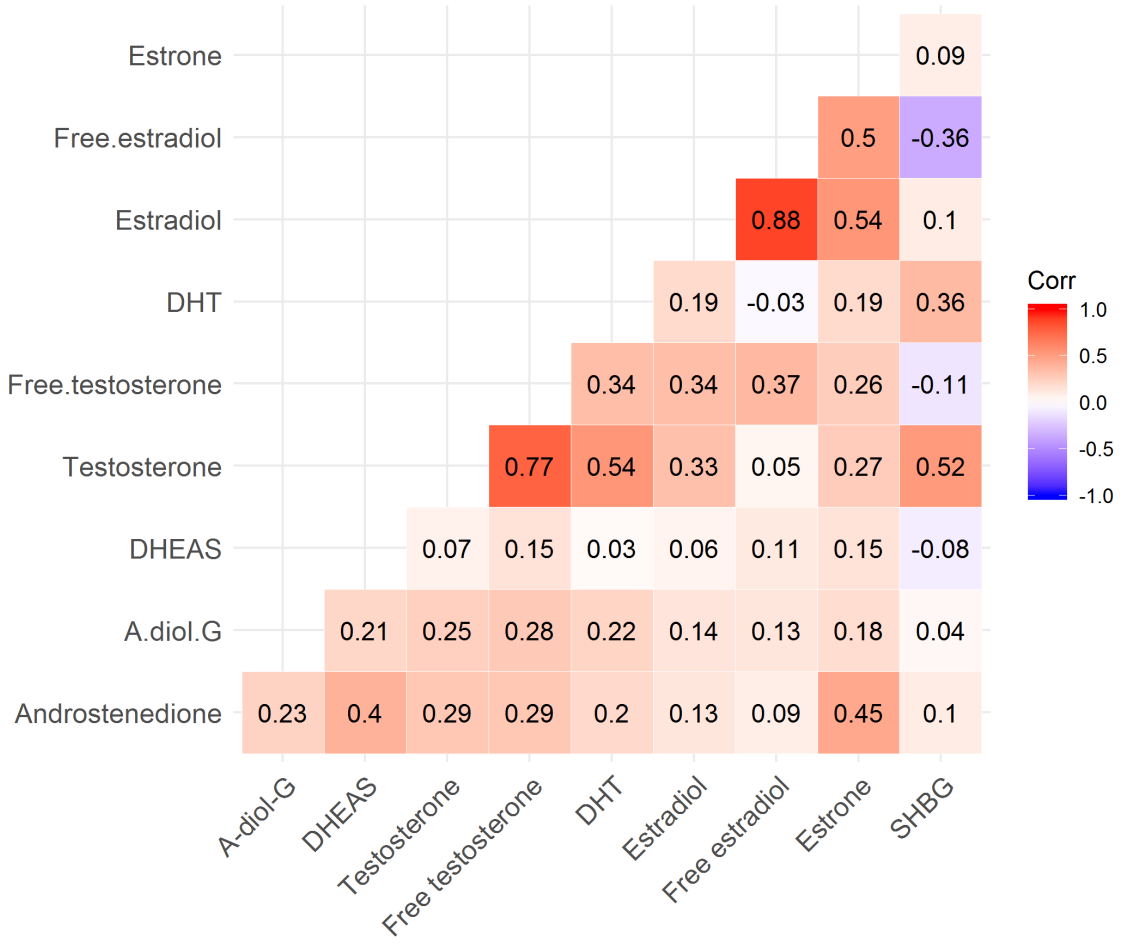


Figure 4.2: Partial correlation coefficients in control subjects between log-transformed concentrations of sex hormones and SHBG.
Values are standardised by study and adjusted by age and BMI.
Abbreviations: A-diol-G=androstanediol glucuronide; BMI=body mass index; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

4.3.2 Age at blood collection

After adjusting for study and BMI, age was associated with concentrations of all hormones, except total testosterone and estradiol (Figure 4.3). Compared to men aged 50-54 years, those aged 75+ years had circulating concentrations of DHEAS, free testosterone, androstenedione, A-diol-G and free estradiol that were 50%, 27%, 26%, 17%, 16% lower, respectively. In contrast, SHBG and DHT concentrations were 39% and 16% higher, respectively. There was evidence of heterogeneity between studies in the associations of free testosterone, DHT, estradiol, free estradiol, estrone and SHBG.

Although there was no evidence of a linear association between age and testosterone concentration, there was evidence of heterogeneity in the association of testosterone concentration with age by time of day of blood collection (F=3.00; P=0.0005). For men whose blood was collected in the morning, younger men (aged 50-54 years) had 10% higher circulating concentrations of total testosterone than those aged 75+ years, while age was not associated with total testosterone concentrations in men whose blood was collected in the afternoon (Figure 4.4).

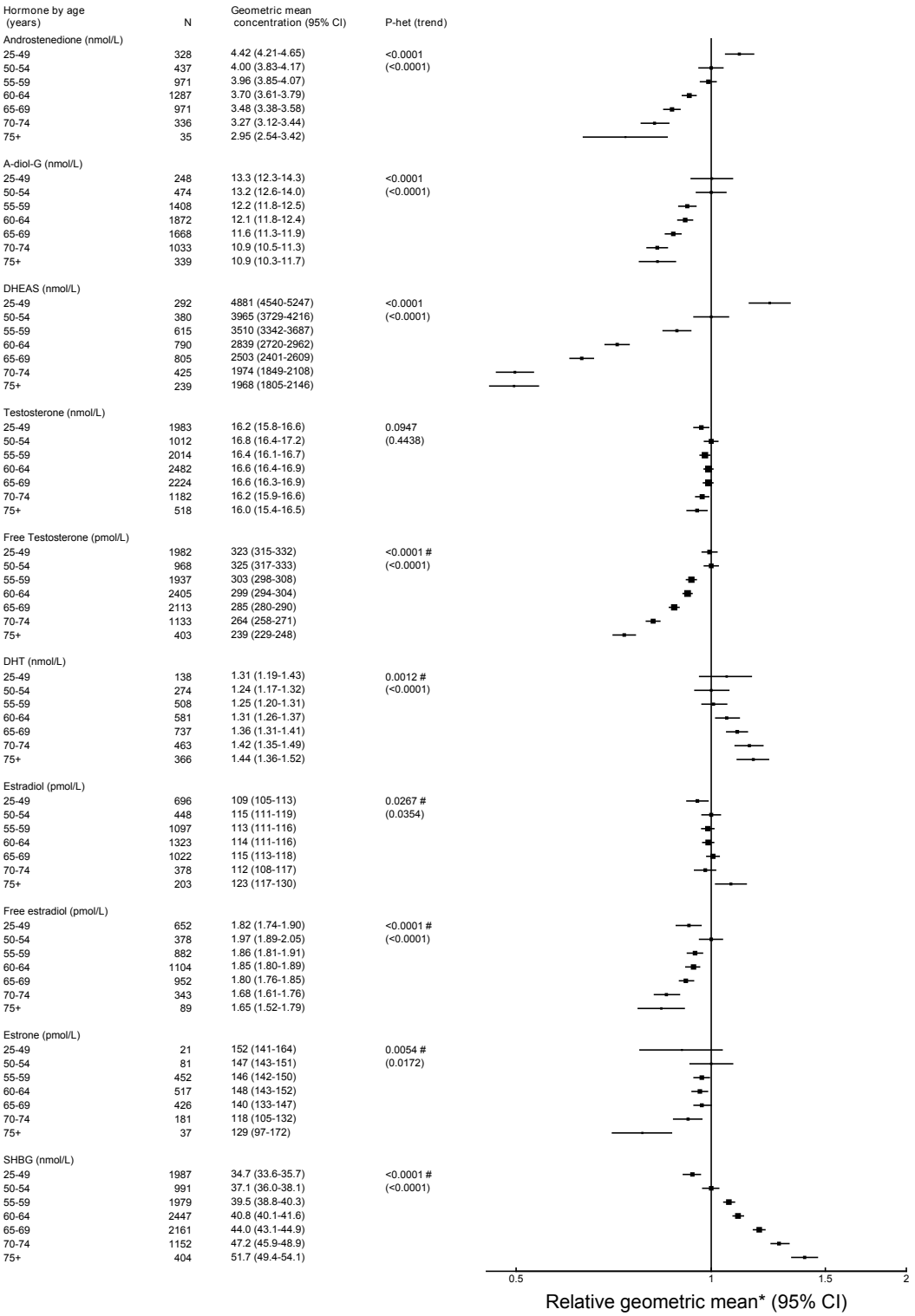


Figure 4.3: Relative geometric mean concentrations* of male sex hormones by age category.
*relative to 50-54 years, adjusted for study and BMI.
#significant interaction with study ($p<0.01$).
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

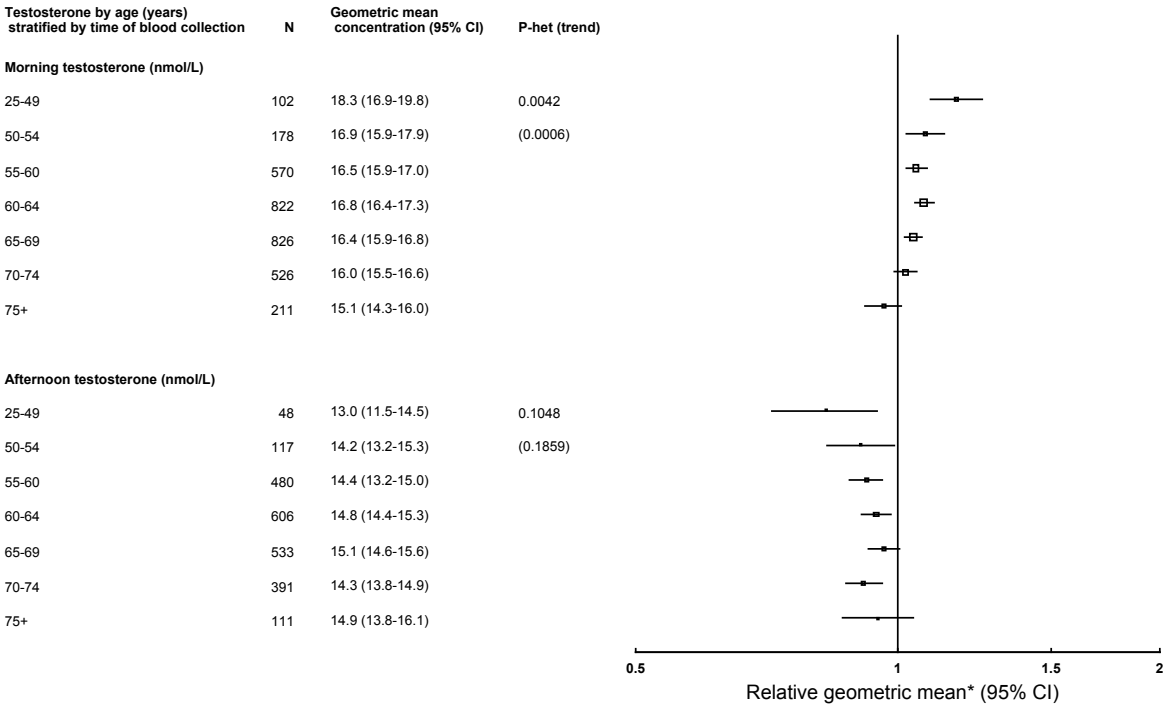


Figure 4.4: Relative geometric mean concentrations* of testosterone by age category, stratified by time of blood collection (morning or afternoon).
*Relative to the overall geometric mean of testosterone, adjusted for study and BMI.

Test of heterogeneity between morning and afternoon blood collection: F=3.00; p=0.0005.
Abbreviations: BMI=body mass index; CI=confidence interval.

4.3.3 Adiposity

After adjusting for study and age, BMI was strongly associated with all sex hormones and SHBG concentrations (Figure 4.5). Compared to men with a BMI 20.0-22.4 kg/m^2, those with a BMI 37.5+ kg/m^2 had concentrations of DHT, testosterone, SHBG, free testosterone, androstenedione and DHEAS that were 45%, 40%, 35%, 26%, 16% and 12% lower, respectively. In contrast, concentrations of free estradiol, A-diol-G, estradiol and estrone were 53%, 38%, 25% and 22% higher, respectively. These associations remained broadly similar when using waist circumference or WHR (Figures 4.6 and 4.7). Differences in geometric mean hormone concentrations by waist circumference and WHR were attenuated following additional adjustment for BMI, although testosterone and SHBG remained associated with both waist circumference and WHR (Figures 4.6 and 4.7).

There was evidence of heterogeneity in the associations of testosterone, free testosterone and SHBG with BMI by self-report and measured methodologies. This appeared to be due to differences in the magnitude rather than the direction of associations (data not shown).

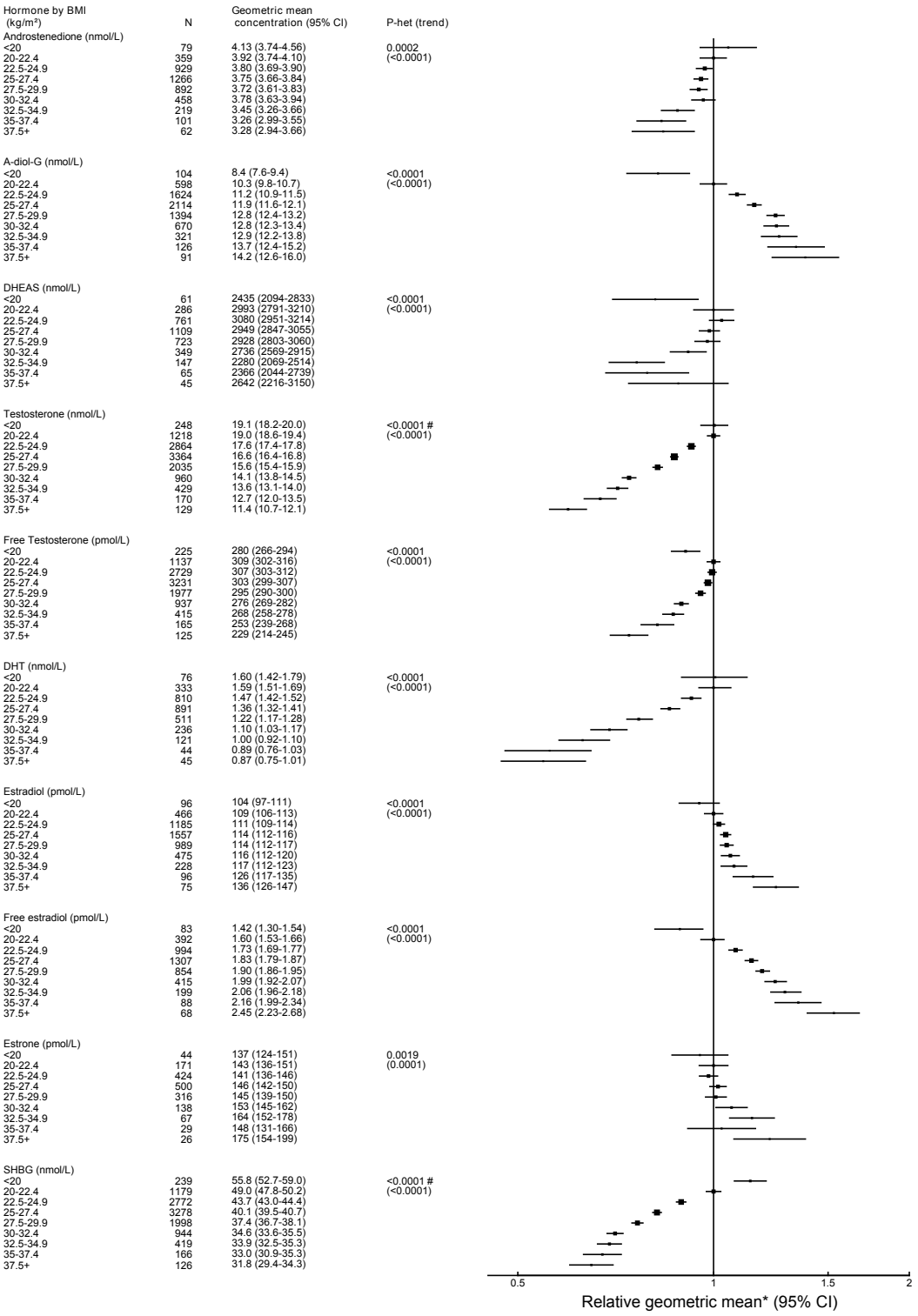


Figure 4.5: Relative geometric mean concentrations* of male sex hormones by BMI category.
*relative to BMI 20.0-22.4 kg/m^2, adjusted for study and age.
#significant interaction with study (p<0.01).
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

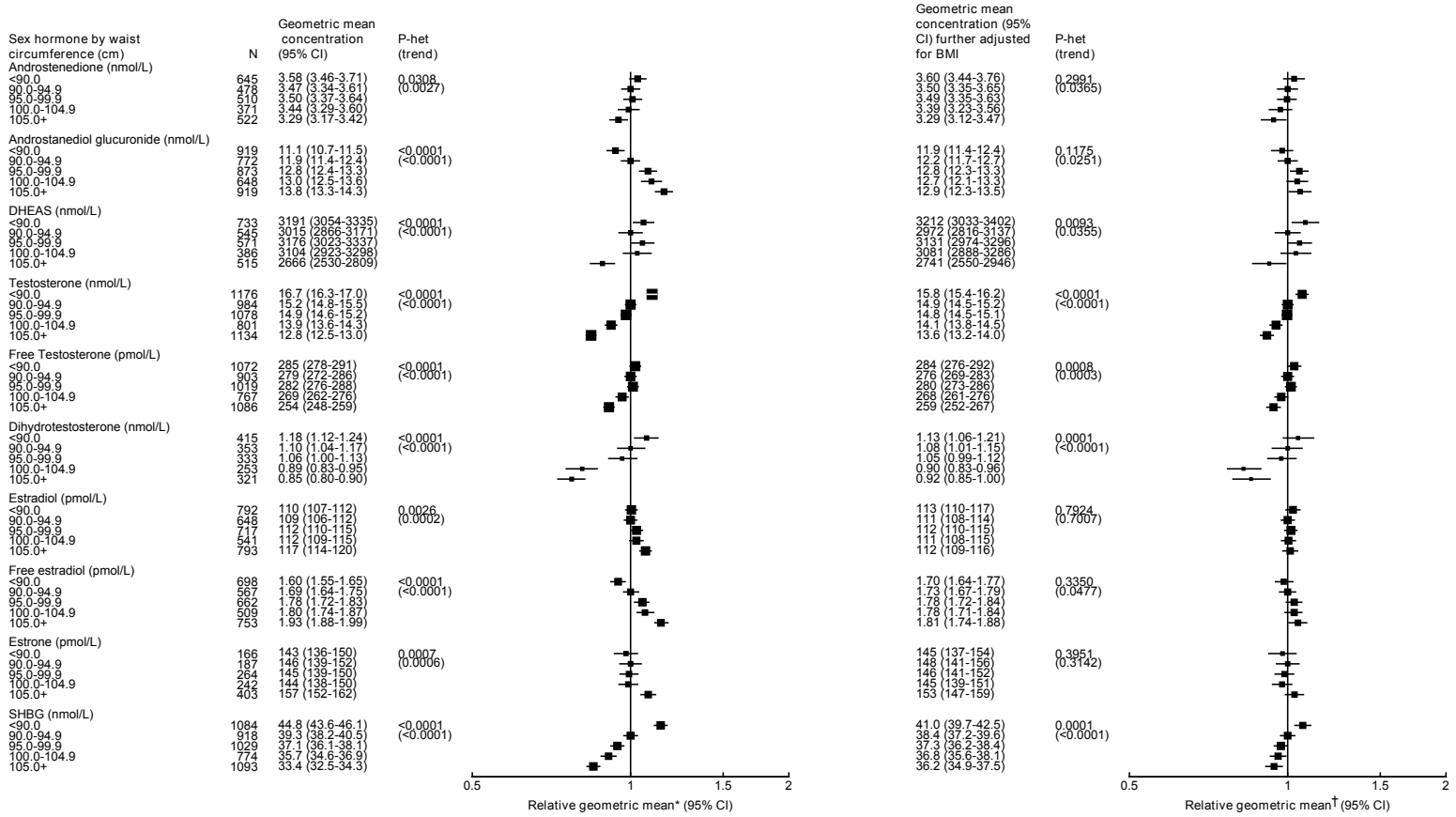


Figure 4.6: Relative geometric mean concentrations* of male sex hormones by waist circumference.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
*relative to 90.0-94.9 cm, adjusted for study and age.
† relative to 90.0-94.9 cm, adjusted for study, age and BMI.
Abbreviations: A-diol-G=androstanediol glucuronide; BMI=body mass index; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

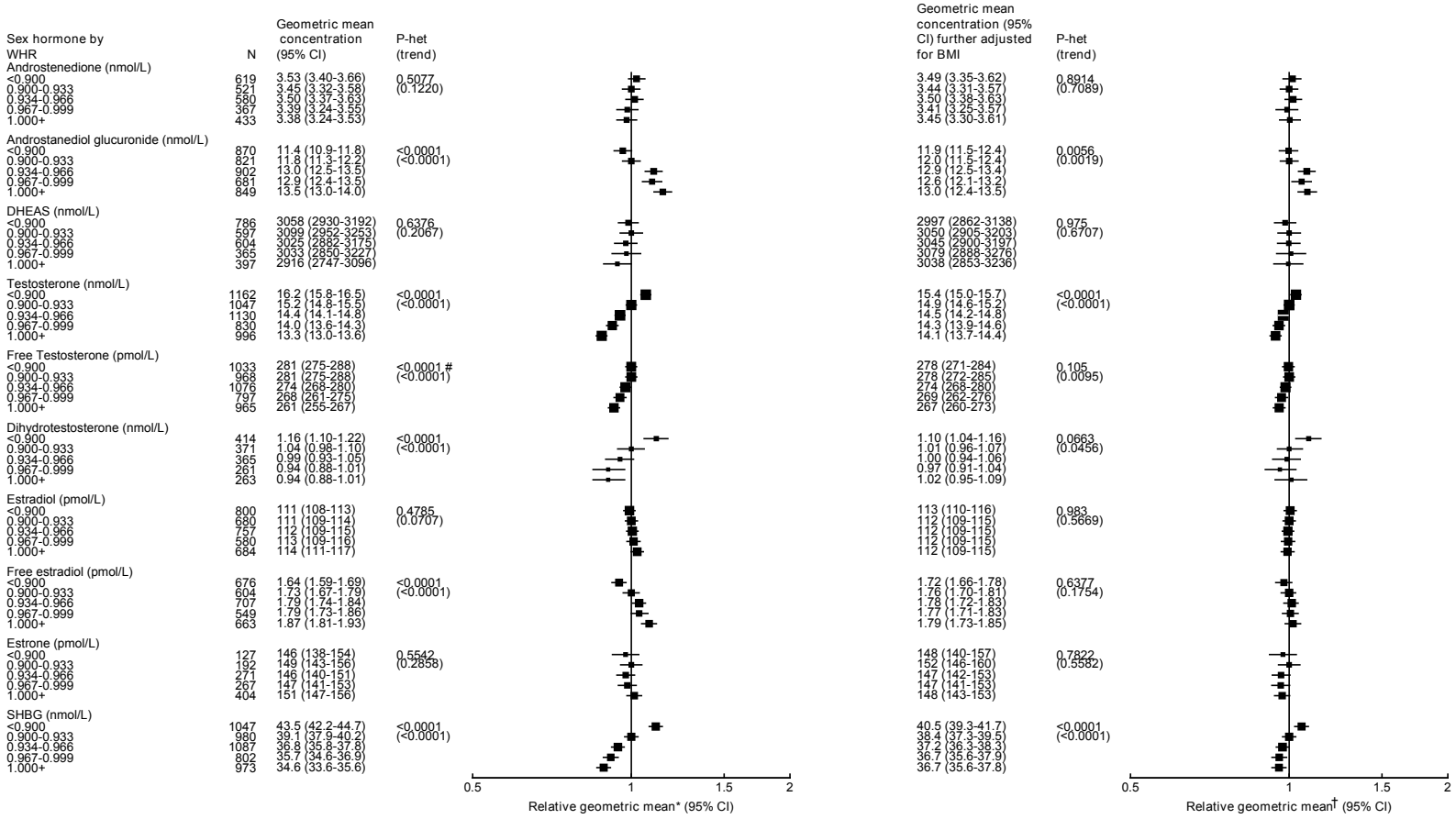


Figure 4.7: Relative geometric mean concentrations* of male sex hormones by WHR.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
*relative to 0.900-0.933, adjusted for study and age.
† relative to 0.900-0.933, adjusted for study, age and BMI.
Abbreviations: A-diol-G=androstanediol glucuronide; BMI=body mass index; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin; WHR=waist-to-hip ratio.

4.3.4 Height

After adjusting for study, age and BMI, height was associated with concentrations of several androgens and SHBG, but not estrogens (Figure 4.8). Compared to men with a height of 160.0-164.9 cm, those in the tallest category (190.0+ cm) had concentrations of androstenedione, testosterone, free testosterone and SHBG that were 17%, 11%, 7%, and 6% lower, respectively, whilst A-diol-G concentration was 4% higher.

4.3.5 Smoking

After adjusting for study, age, BMI and alcohol consumption, smoking status was positively associated with concentrations of androgens and SHBG, but not estrogens (Figure 4.9). Compared with never smokers, heavy smokers (30+ cigarettes/day) had concentrations of androstenedione, SHBG and testosterone that were 24%, 11% and 6% higher, respectively. The associations between smoking intensity and these androgens and SHBG remained when the analyses were restricted to current smokers only.

4.3.6 Alcohol

After adjusting for study, age, BMI and smoking, alcohol consumption was associated with higher concentrations of several androgens, but was not associated with estrogens (Figure 4.10). In particular, compared with low alcohol consumers (1-9 g ethanol/day), those who drank 70+ g ethanol/day had a 13% higher concentration of DHEAS. High alcohol consumption was also moderately associated with higher concentrations of androstenedione and A-diol-G (9% and 7% higher, respectively).

Non-drinkers had 14% and 10% lower concentrations of DHEAS and DHT, respectively in comparison with men who reported drinking 1-9 g ethanol per day.

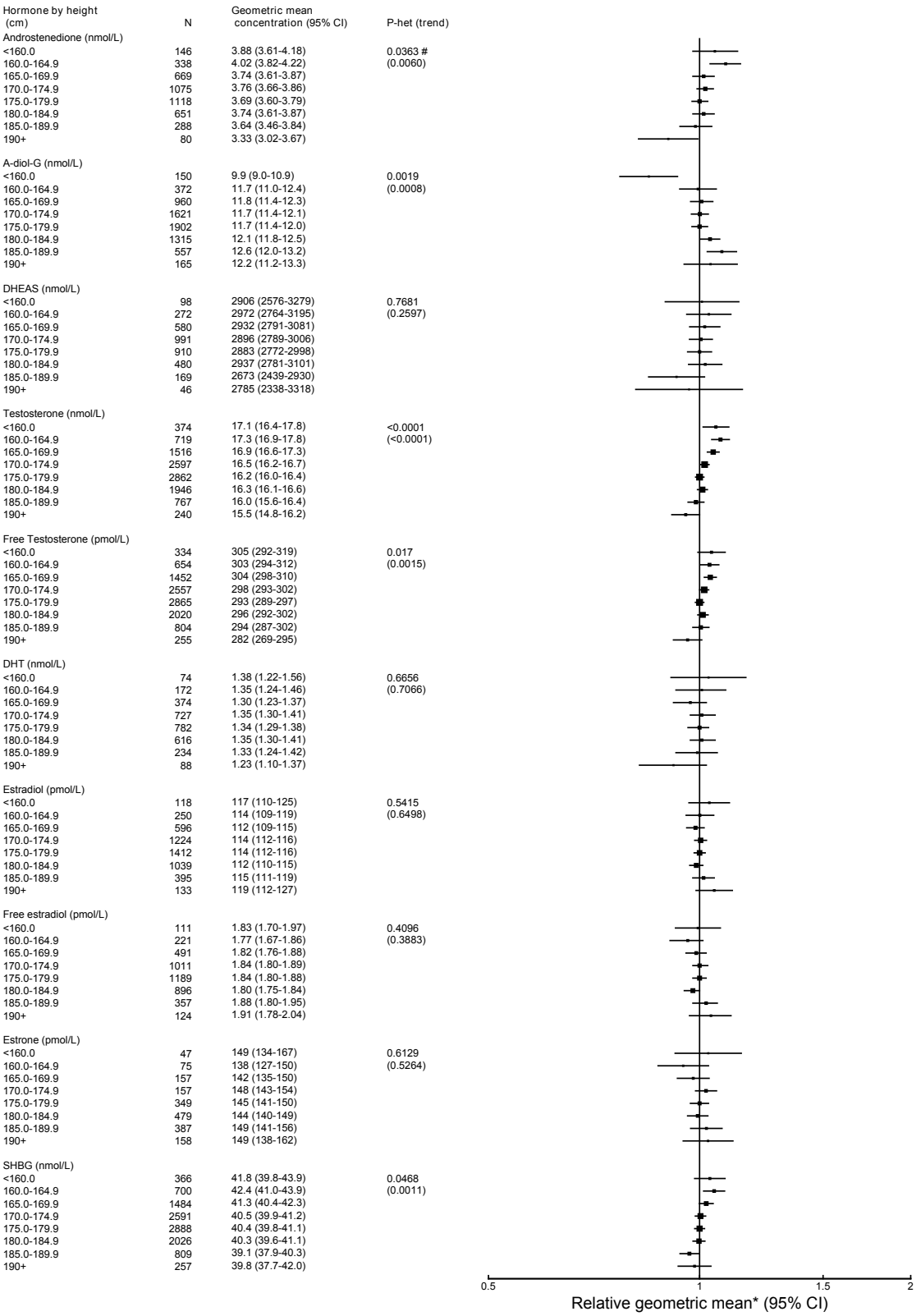


Figure 4.8: Relative geometric mean concentrations* of male sex hormones by height, adjusted for study, age and BMI.
*relative to 160.0-164.9cm.
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

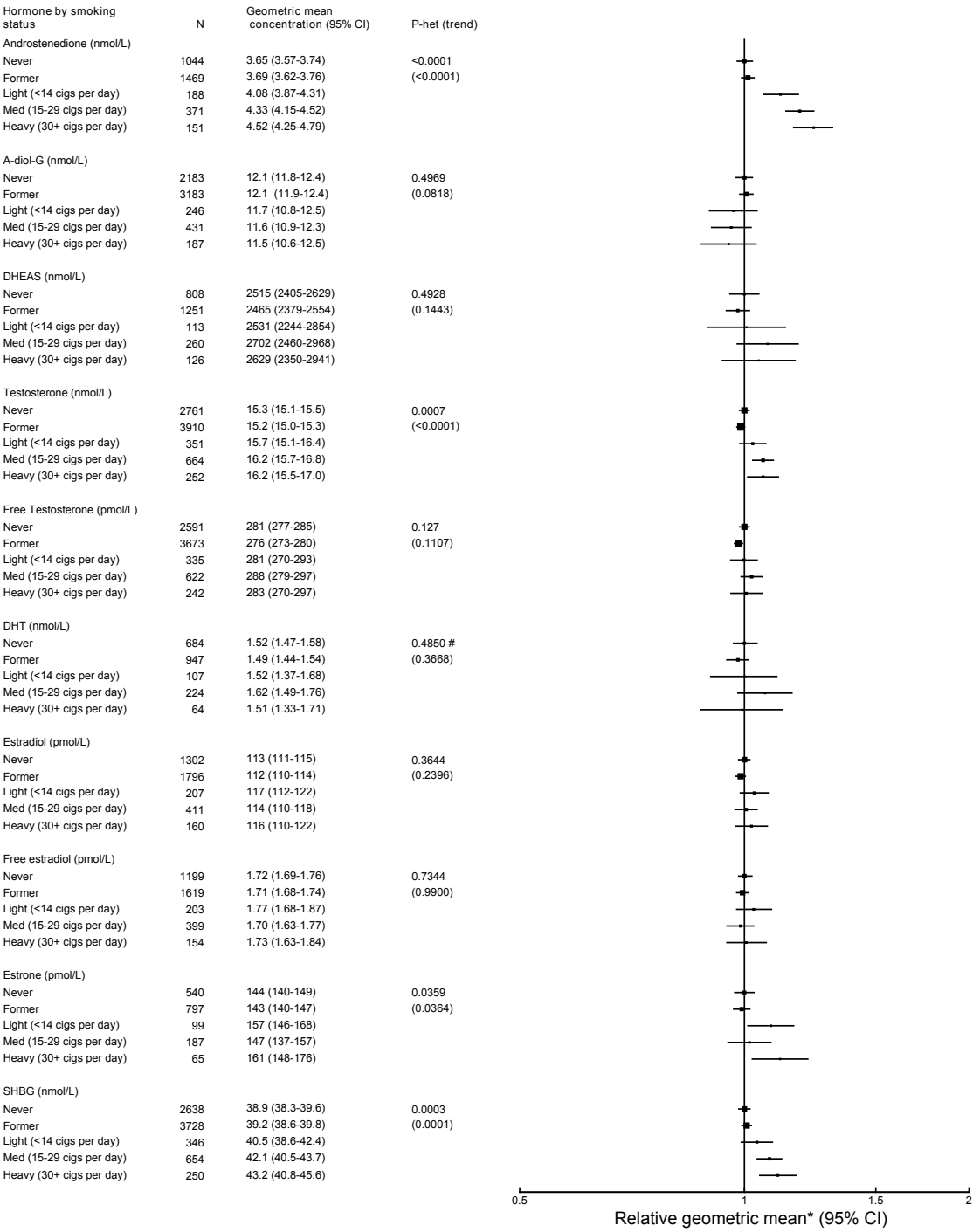


Figure 4.9: Relative geometric mean concentrations* of male sex hormones by smoking status, adjusted for study, age, BMI and alcohol consumption.
*relative to never smokers.
#significant interaction with study ($p<0.01$).
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

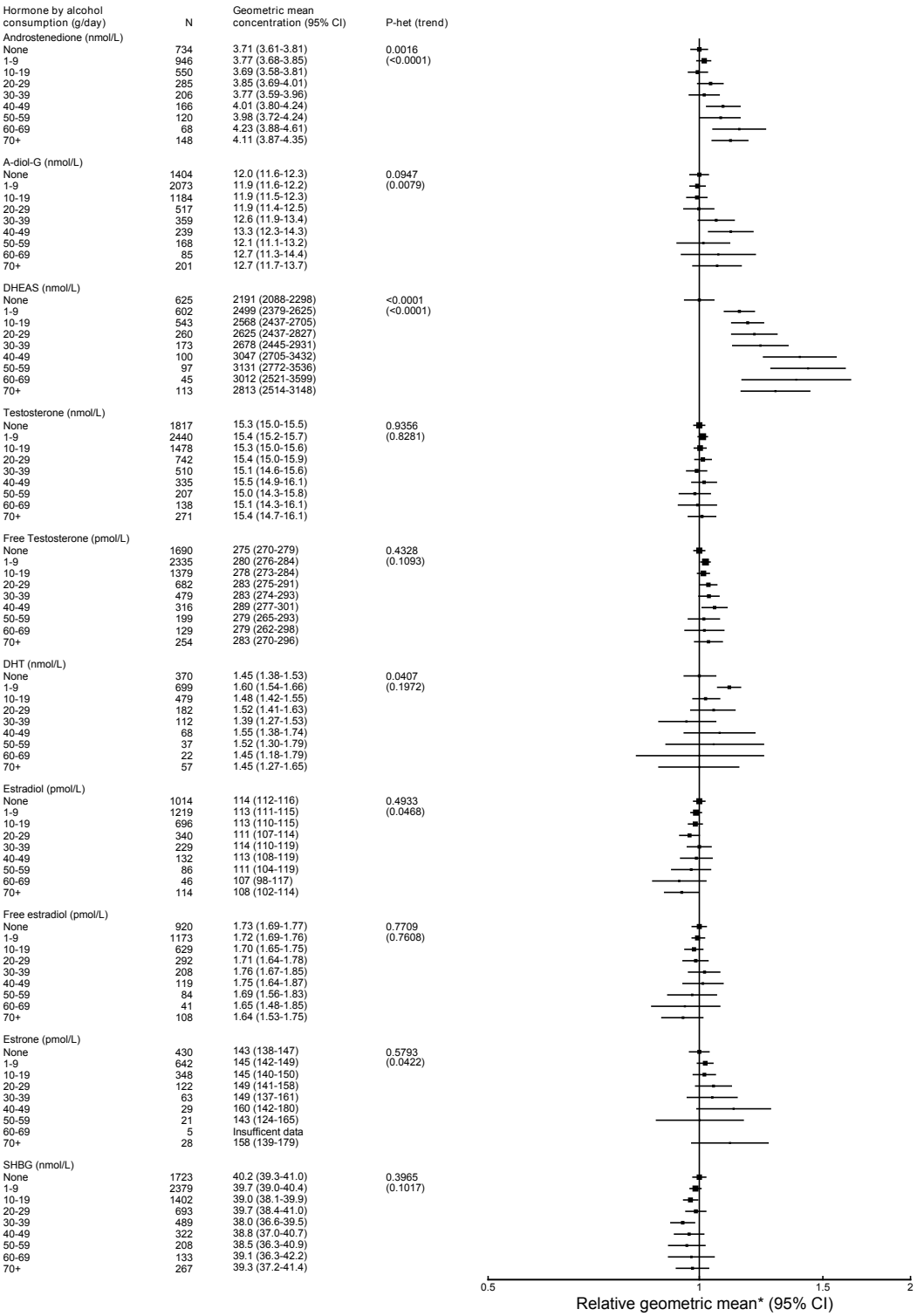


Figure 4.10: Relative geometric mean concentrations* of male sex hormones by alcohol consumption, adjusted for study, age, BMI and smoking status.
*relative to men with low alcohol consumption (1-9 g ethanol/day).
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

4.3.7 Ethnic/racial group

After adjusting for study, age and BMI, racial/ethnic group was not associated with sex hormone concentrations, with the exception of A-diol-G and estrogens (Figure 4.11). Compared with non-Hispanic white men, A-diol-G concentration was 25% lower in East Asian men and 12% higher in Hispanic/Latino men, while estrone, estradiol and free estradiol concentrations were 15%, 14% and 12% higher in African American men than non-Hispanic whites.

4.3.8 Time of blood collection

After adjusting for study, age and BMI, time of blood collection was associated with several sex hormone concentrations (Figure 4.12). Compared with men whose blood was collected before 09:00, participants whose blood was collected at 15:00 or later had concentrations of androstenedione, testosterone, free testosterone and A-diol-G that were 19%, 16%, 15% and 8% lower, respectively.

4.3.9 Sociodemographic factors and health

After adjusting for study, age and BMI, sociodemographic factors had small magnitudes of associations with sex hormones. In comparison with men with no secondary/high school qualifications, men with a university education had concentrations of SHBG, androstenedione and testosterone that were 7%, 6% and 5% lower, respectively (Figure 4.13).

In comparison with men who were married/cohabiting, men who were not married/cohabiting had concentrations of estrone, androstenedione, free estradiol and estradiol and that were 12%, 7%, 6% and 6% higher, respectively (Figure 4.14). Family history of prostate cancer was not associated with circulating sex steroid hormone concentrations (Figure 4.15).

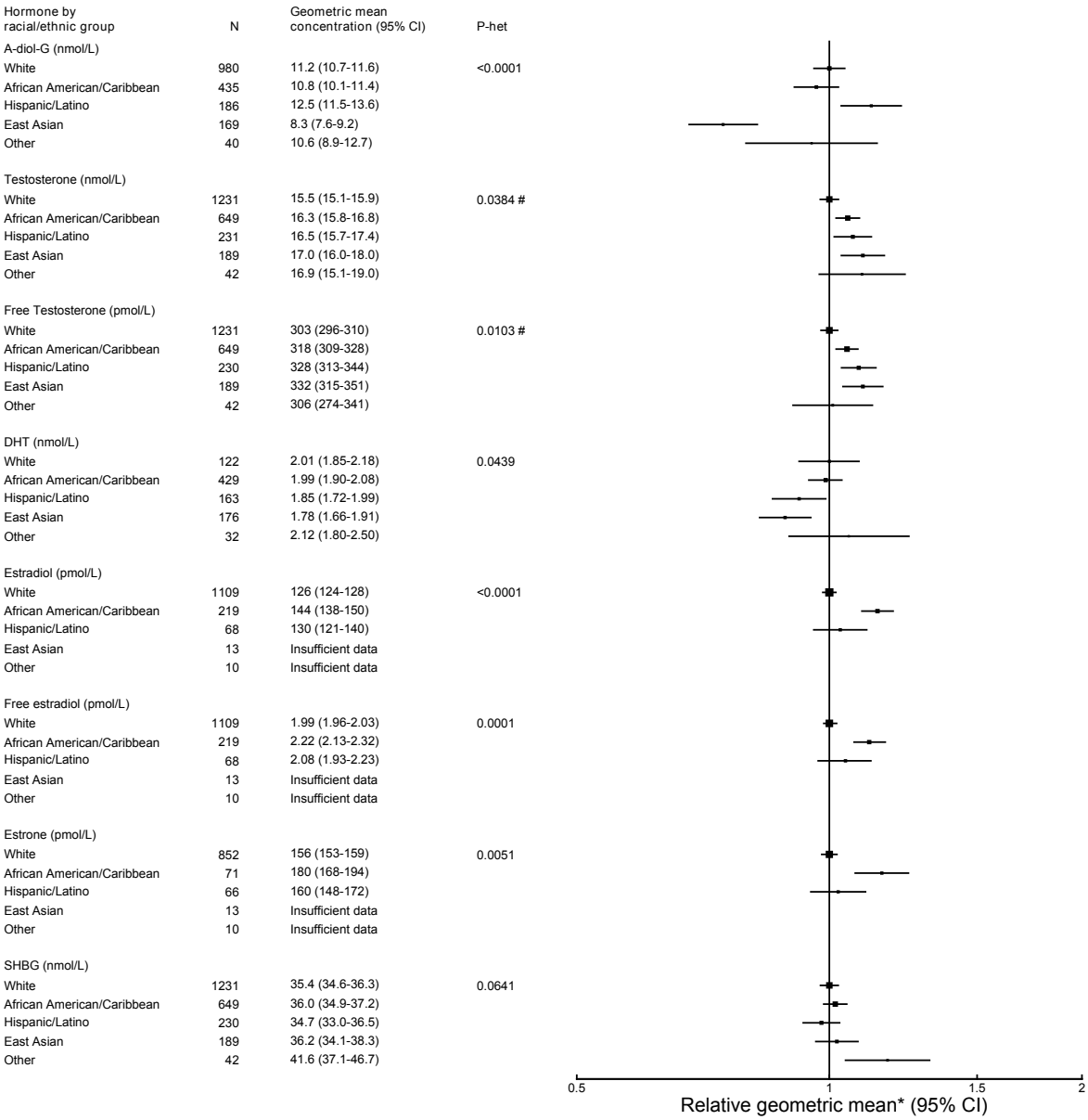


Figure 4.11: Relative geometric mean concentrations* of male sex hormones by racial/ethnic group, adjusted for study, age and BMI.
*relative to non-Hispanic whites.
#significant interaction with study ($p<0.01$).
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

Hormone by time of blood collection	N	Geometric mean concentration (95% CI)	P-het
Androstenedione (nmol/L)			
Before 9:00	526	4.34(4.19-4.50)	<0.0001
9:00-11:59	1051	4.06 (3.96-4.16)	
12:00-14:59	784	4.07 (3.95-4.19)	
After 15:00	283	3.51 (3.34-3.69)	
A-diol-G (nmol/L)			
Before 9:00	939	12.5 (12.0 -13.0)	0.0021
9:00-11:59	1943	11.5 (11.3-11.8)	
12:00-14:59	1353	11.4 (11.1-11.8)	
After 15:00	714	11.5 (11.0-12.0)	
DHEAS (nmol/L)			
Before 9:00	250	2212 (2024-2417)	0.0158
9:00-11:59	719	2232 (2136-2332)	
12:00-14:59	320	2339 (2178-2511)	
After 15:00	233	2606 (2401-2830)	
Testosterone (nmol/L)			
Before 9:00	1318	16.3 (16.0-16.7)	<0.0001 #
9:00-11:59	2440	15.7 (15.5-16.0)	
12:00-14:59	1566	15.4 (15.1-15.7)	
After 15:00	725	13.8 (13.4-14.3)	
Free Testosterone (pmol/L)			
Before 9:00	1141	301 (295-308)	<0.0001 #
9:00-11:59	2243	292 (287-296)	
12:00-14:59	1502	282 (277-287)	
After 15:00	700	254 (247-262)	
DHT (nmol/L)			
Before 9:00	776	1.69 (1.64-1.75)	0.0533
9:00-11:59	826	1.67 (1.62-1.72)	
12:00-14:59	284	1.63 (1.54-1.72)	
After 15:00	54	1.42 (1.25-1.61)	
Estradiol (pmol/L)			
Before 9:00	337	115 (111-119)	0.2325
9:00-11:59	811	113 (111-115)	
12:00-14:59	448	114 (110-117)	
After 15:00	336	110 (106-113)	
Free estradiol (pmol/L)			
Before 9:00	194	1.86 (1.78-1.94)	0.3031
9:00-11:59	653	1.86 (1.82-1.90)	
12:00-14:59	448	1.84 (1.79-1.89)	
After 15:00	336	1.79 (1.73-1.85)	
Estrone (pmol/L)			
Before 9:00	284	152 (146-158)	0.1473
9:00-11:59	698	145 (142-149)	
12:00-14:59	401	146 (141-151)	
After 15:00	293	142 (137-148)	
SHBG (nmol/L)			
Before 9:00	1168	39.6 (38.7-40.6)	0.3938
9:00-11:59	2290	39.2 (38.5-39.8)	
12:00-14:59	1553	39.9 (39.1-40.7)	
After 15:00	719	39.0 (37.8-40.2)	

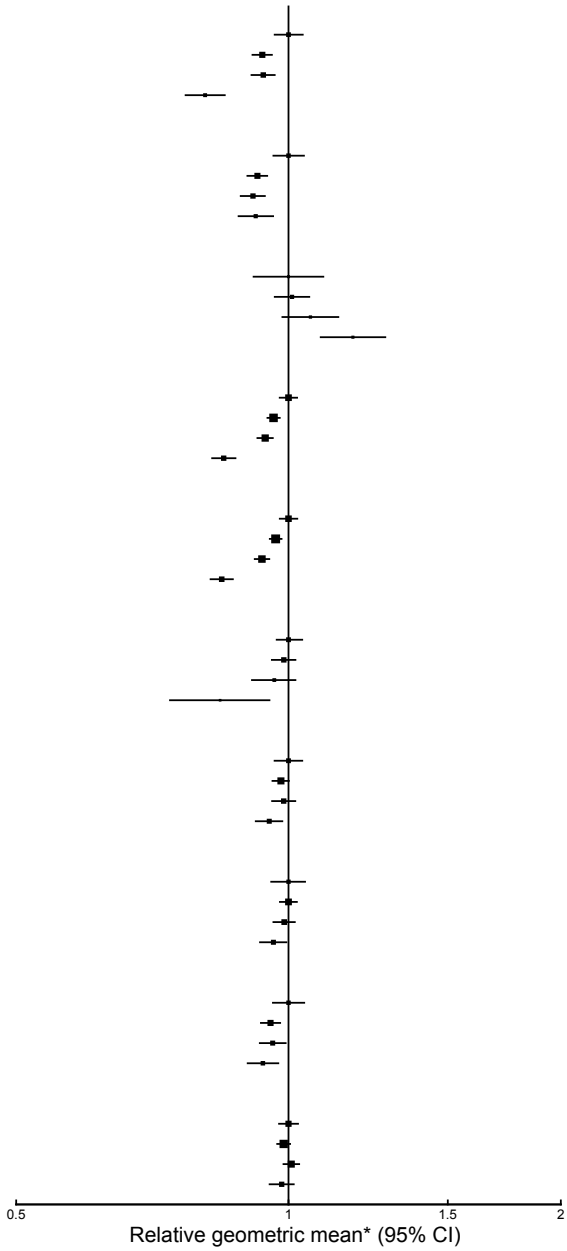


Figure 4.12: Relative geometric mean concentrations* of male sex hormones by time of blood collection.
*relative to before 09:00, adjusted for study, age and BMI.
#significant interaction with study P<0.01.
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

Hormone by education status	N	Geometric mean concentration (95% CI)	P-het (trend)
Androstenedione (nmol/L)			
< Secondary/ High School	854	3.79 (3.68-3.90)	0.011
Secondary/ High School	1241	3.68 (3.59-3.77)	(0.0050)
College	726	3.74 (3.62-3.86)	
University and above	1134	3.56 (3.47-3.65)	
A-diol-G (nmol/L)			
< Secondary/ High School	1349	11.7 (11.3-12.1)	0.8286
Secondary/ High School	1748	11.9 (11.6-12.2)	(0.4369)
College	1147	12.0 (11.6-12.4)	
University and above	1564	12.0 (11.6-12.3)	
DHEAS (nmol/L)			
< Secondary/ High School	841	2918 (2791-3052)	0.4556
Secondary/ High School	875	2816 (2701-2936)	(0.8732)
College	961	2951 (2836-3072)	
University and above	773	2882 (2757-3013)	
Testosterone (nmol/L)			
< Secondary/ High School	1508	17.3 (16.9-17.7)	0.001
Secondary/ High School	2154	16.7 (16.3-17.1)	(0.0028)
College	1724	17.0 (16.6-17.5)	
University and above	2115	16.4 (16.0-16.8)	
Free Testosterone (pmol/L)			
< Secondary/ High School	1444	276 (270-283)	0.5447
Secondary/ High School	2025	275 (269-281)	(0.7760)
College	1543	279 (272-286)	
University and above	2027	274 (268-280)	
DHT (nmol/L)			
< Secondary/ High School	408	1.34 (1.27-1.42)	0.0684
Secondary/ High School	558	1.36 (1.31-1.42)	(0.4219)
College	630	1.45 (1.40-1.51)	
University and above	634	1.37 (1.32-1.43)	
Estradiol (pmol/L)			
< Secondary/ High School	590	118 (114-121)	0.0555
Secondary/ High School	1376	117 (115-119)	(0.1160)
College	834	119 (117-122)	
University and above	1130	115 (112-117)	
Free estradiol (pmol/L)			
< Secondary/ High School	588	1.90 (1.84-1.96)	0.1678
Secondary/ High School	1323	1.93 (1.90-1.97)	(0.9081)
College	658	1.97 (1.92-2.03)	
University and above	1063	1.91 (1.87-1.95)	
Estrone (pmol/L)			
< Secondary/ High School	223	149 (140-158)	0.7711
Secondary/ High School	560	154 (150-158)	(0.7472)
College	160	156 (148-164)	
University and above	396	153 (148-159)	
SHBG (nmol/L)			
< Secondary/ High School	1495	48.3 (47.1-49.4)	<0.0001
Secondary/ High School	2104	46.4 (45.3-47.4)	(<0.0001)
College	1546	47.1 (45.9-48.3)	
University and above	2046	45.0 (43.9-46.0)	

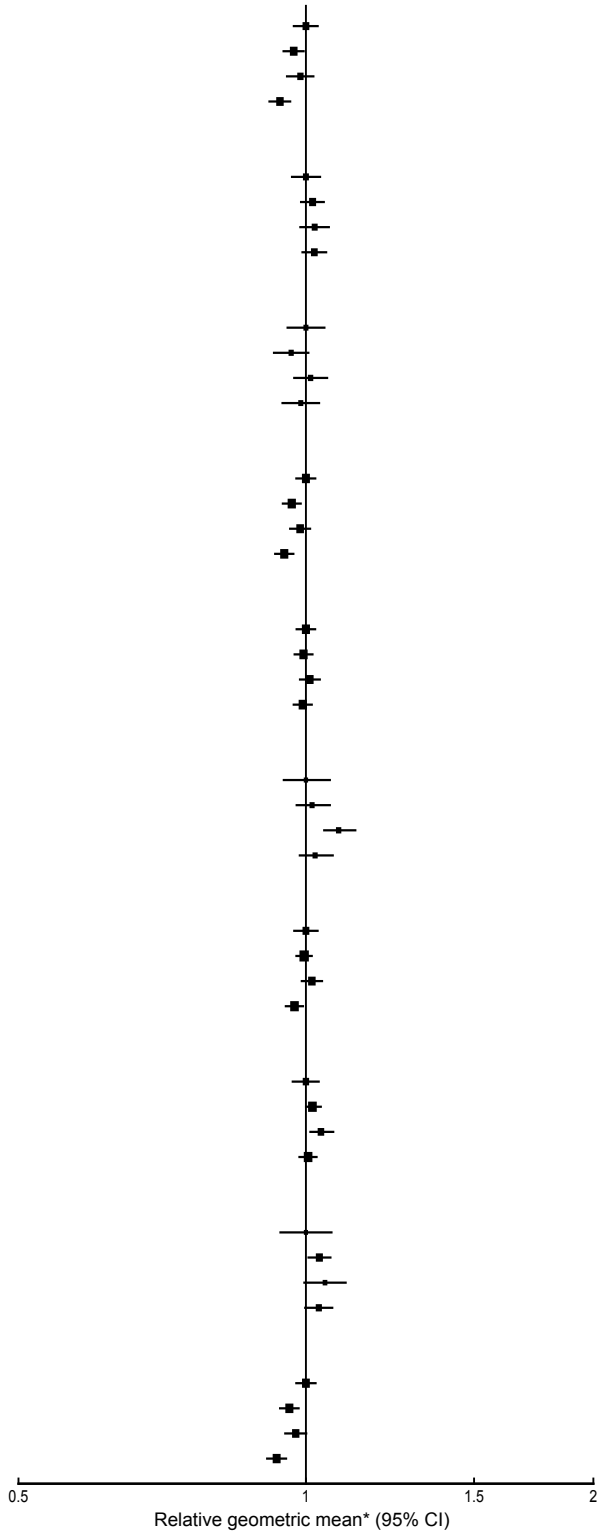


Figure 4.13: Relative geometric mean concentrations* of male sex hormones by education status, adjusted for study, age and BMI.
*relative to < secondary/high school.
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

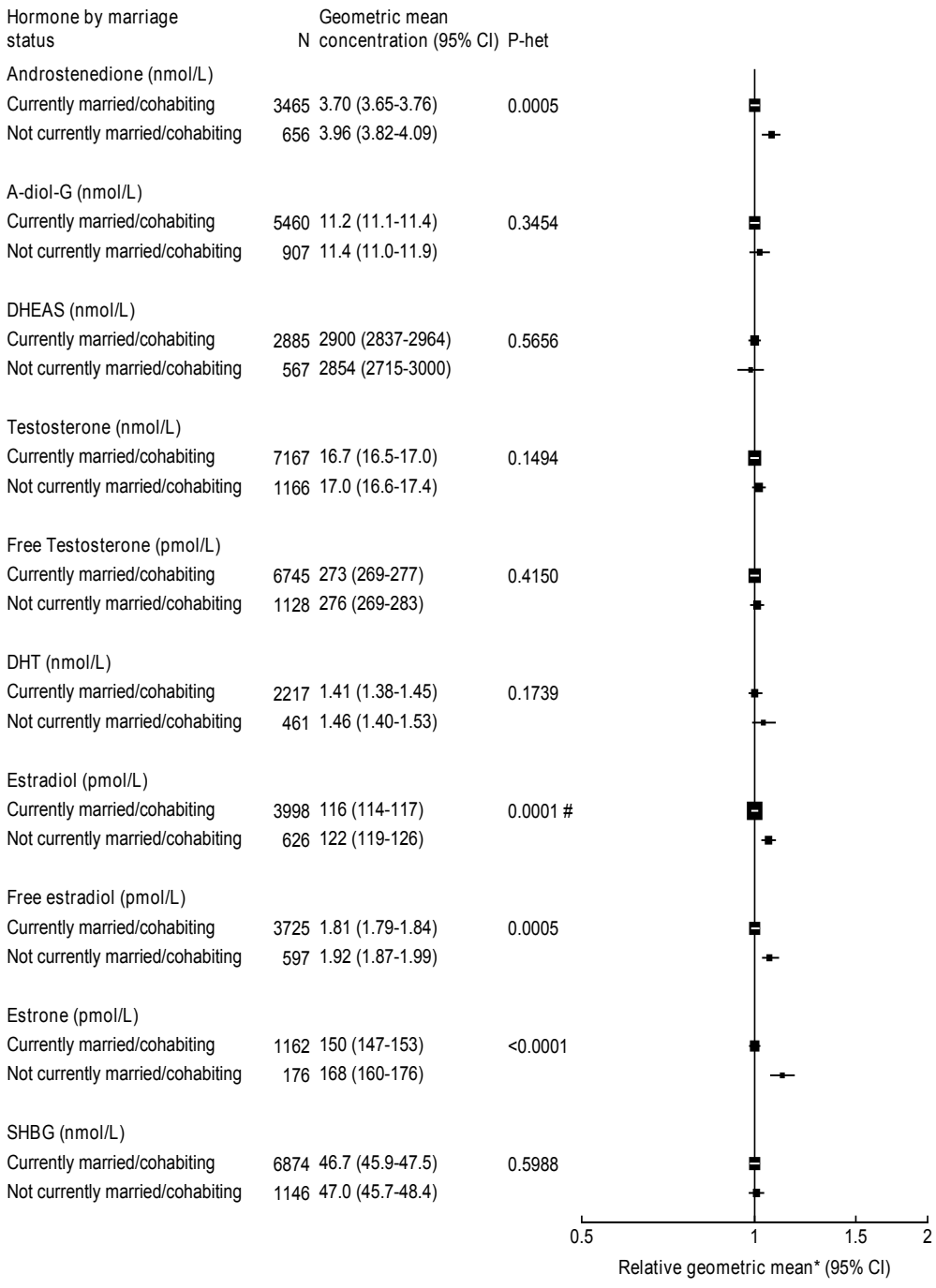


Figure 4.14: Relative geometric mean concentrations* of male sex hormones by marriage status, adjusted for study, age and BMI.
*relative to not married/cohabiting.
#significant interaction with study ($p<0.01$).
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

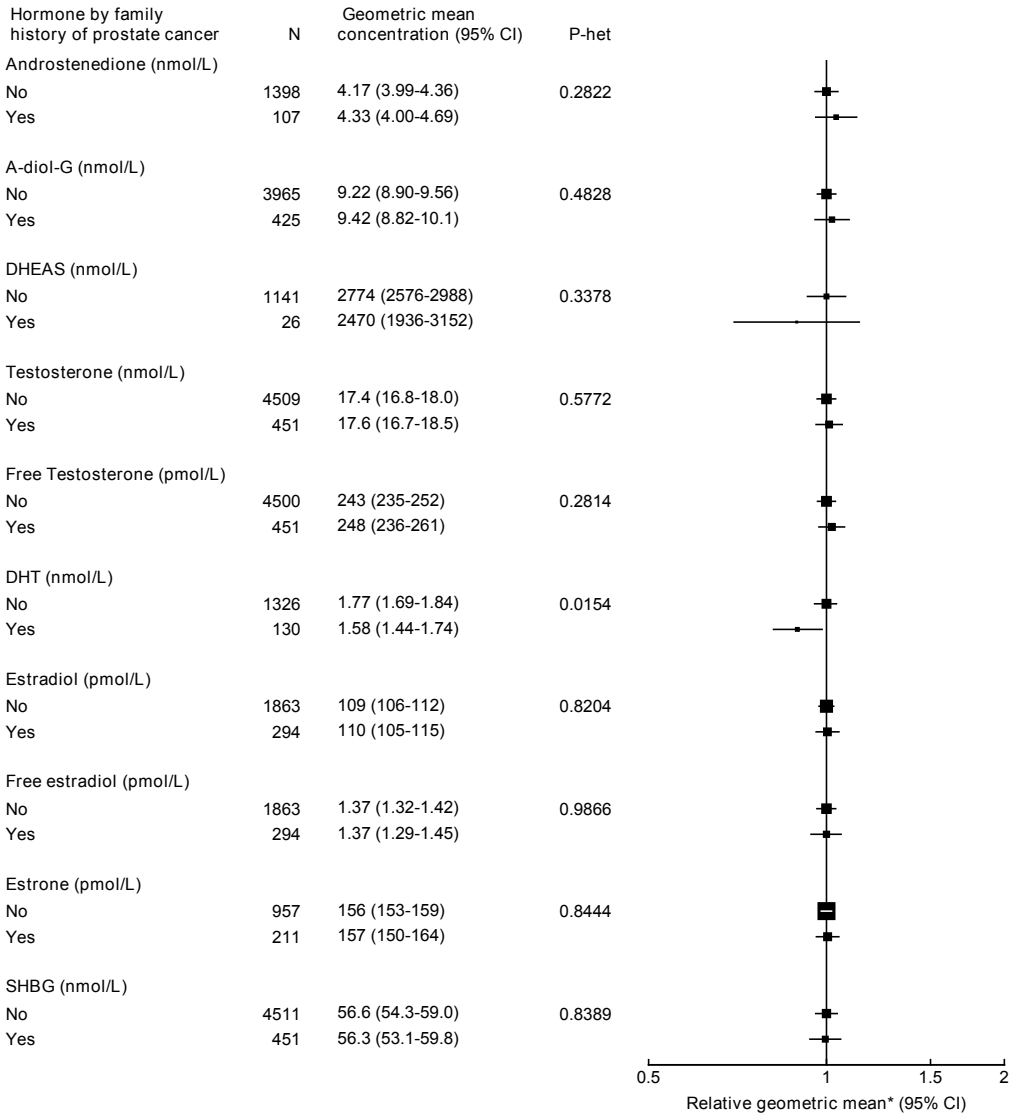


Figure 4.15: Relative geometric mean concentrations* of male sex hormones by family history of prostate cancer, adjusted for study, age and BMI.
*relative to no family history of prostate cancer.
Abbreviations: A-diol-G=androstanediol glucuronide; CI=confidence interval; DHEAS=dehydroepiandrosterone sulfate; DHT=dihydrotestosterone; SHBG=sex hormone binding globulin.

4.3.10 Heterogeneity by study and assay

Some heterogeneity between studies in the associations anthropometric and sociodemographic factors with sex hormone concentrations was observed (see Figures 4.3-4.15). Much of the observed heterogeneity was due to differences in the magnitude rather than the direction of associations (data not shown). Some of this heterogeneity may be explained by differences in assay type. There was heterogeneity by assay type in the associations of DHT and estradiol concentrations with age; there was no increase in DHT with age in studies that used assays without an extraction step prior to assay, whilst age was associated with higher DHT concentration in studies with an extraction step. For estradiol, the positive association with age was stronger in studies that used an extraction step than in those that did not.

4.3.11 Sensitivity analyses

Stratification by age, BMI and time of day did not materially affect the results (data not shown), with the exception of the association of total testosterone with age (Figure 4.4). The results also remained broadly unchanged after: i) restricting the dataset to non-Hispanic white men; ii) excluding within study hormone outliers; iii) following further adjustment for height, smoking and alcohol; iv) using alternative calculations of free testosterone and free estradiol.

4.4 Discussion

This international collaboration has brought together and analysed data from 12,300 men and examined the associations of various anthropometric, behavioural and sociodemographic factors with circulating sex hormone and SHBG concentrations. These findings suggest that age, body composition and, to a lesser extent, smoking status and alcohol consumption, may be important determinants of circulating sex hormone concentrations.

4.4.1 Correlations between hormones

As expected, sex hormones were strongly correlated with other hormones close to them on the sex steroid pathway (see Chapter 2). Testosterone is primarily secreted by the testes. The adrenal glands are another source of androgens, including DHEAS and androstenedione, both of which can be converted to testosterone and/or DHT in the peripheral tissues. Androgens can also be converted into estrogens, primarily in the testes and adipose tissue. <3% of circulating testosterone and estradiol is "free" and therefore able to pass out of the bloodstream and into the peripheral tissues, the remainder is primarily bound to either SHBG or albumin [160].

4.4.2 Age

Age was strongly associated with circulating SHBG and androgens, except for total testosterone concentration. Other studies have generally found that total testosterone declines with age [300, 301, 308–311], although the observed decline may be partly due to confounding by obesity and co-morbidities [312–316]. The current analysis found no evidence of an overall association of total testosterone with age. However, stratification by time of day showed a linear decline with age only in men whose blood was collected in the morning; circulating testosterone concentrations are highest in the morning and decline throughout the day, and this diurnal rhythm may be blunted as men age [317, 318]. The marked positive associations of SHBG concentrations with age may be partly related to a lower insulin-like growth factor-I (IGF-I) concentrations in older individuals [319], as IGFs may downregulate hepatic

SHBG production [162]. Adrenal sensitivity to adrenocorticotrophic hormone has also been shown to decline with age [320], which results in a substantially reduced production of DHEAS and, to a lesser extent, androstenedione [321, 322]. Despite this age-associated decline in androstenedione, DHT concentrations were higher in older men, which might be due to the reduced conversion of DHT to A-diol-G [322].

4.4.3 Adiposity

BMI was strongly associated with concentrations of all the sex hormones and SHBG. A higher BMI was associated with markedly lower SHBG concentrations, which may be due to high insulin [323], low-grade inflammation and/or increased liver fat [228, 324]. Lower SHBG concentrations may reduce testosterone production via a negative feedback loop. Increasing adipose tissue also leads to higher aromatase activity [325], which converts androgens to estrogens, leading to higher concentrations of circulating estrogens. This, together with the low concentrations of SHBG, results in markedly higher concentrations of free estradiol in obese men. The finding of higher A-diol-G concentrations in obese men suggests an increase in peripheral androgen metabolism with greater adiposity. Obesity leads to dysregulation of the hypothalamic-pituitary-adrenal axis [323, 326] and may therefore result in lower DHEAS and androstenedione production. After additional adjustment for BMI, waist circumference and WHR each remained associated with testosterone and SHBG concentrations. This suggests that visceral fat, or other factors correlated with waist circumference and WHR, may be important predictors of androgen and SHBG concentrations [304, 327].

4.4.4 Height

Testosterone, free testosterone and SHBG were weakly inversely associated with height, and A-diol-G was positively associated with height. Estrogens are involved in the control of growth in height [328] and affect growth hormone secretion [329], there is some evidence to suggest that androgens may also be involved [330]. However, it is unclear to what extent circulating sex hormone concentrations during childhood and adolescence correlate with concentrations later in life.

4.4.5 Smoking

The mechanisms through which tobacco may affect hormone concentrations are complex because cigarette smoke contains a wide range of endocrine disruptors, some of which exert opposing physiological effects [331, 332]. Smoking was associated with higher concentrations of testosterone and SHBG, which resulted in a relatively unchanged free testosterone concentrations [333–335]. Smoking was also associated with higher circulating androstenedione concentrations.

4.4.6 Alcohol

Greater alcohol consumption was associated with moderately higher androgen concentrations, particularly DHEAS, as observed previously [336–338] and to a lesser extent androstenedione, possibly involving increased adrenal secretion and inhibition of hydroxysteroid dehydrogenase [338]. The large magnitude of the association observed with DHEAS concentrations may be related to the long half-life of this hormone [152]. Alcohol consumption was also moderately positively associated with A-diol-G, although findings from other studies are inconsistent [339, 340]. Chronic excessive alcohol consumption is a well-known cause of low testosterone in men [341], but this analysis did not find evidence of an association between alcohol consumption and testosterone concentration, perhaps because the duration of consuming the reported quantities of alcohol is unknown.

4.4.7 Racial/ethnic group

Three studies provided data from men across several ethnic/racial groups and this analysis found that ethnic/racial group was associated with circulating A-diol-G concentration, which was lower in East Asian men compared to all other ethnic/racial groups, as observed previously [86, 342, 343]. Estrone, estradiol and free estradiol concentration were also higher in African Americans, as reported elsewhere [85]. The mechanisms for these differences remain unclear. Further inference regarding the association between ethnicity/race and other hormone concentrations was not possible due to the relatively low representation of non-white men.

4.4.8 Heterogeneity between studies

Heterogeneity between the included studies was present for several associations. Further investigation showed that this was generally caused by differences in the magnitude of the associations rather than the direction. Heterogeneity may have been caused by differences in sample populations, study design, and for testosterone time of day of blood draw.

Additionally, some of the observed heterogeneity may have been attributable to differences in hormone measurement methods. In this analysis, no study used mass spectrometry, often considered the gold standard method of measurement for sex hormones [344]. Instead, most sex hormones were measured using extraction or non-extraction immunoassays. Non-extraction assays generally have a lower sensitivity and specificity for hormone concentrations than extraction assays, and differences in assay performance may explain some of the differences observed in hormone concentrations.

4.4.9 Calculation of free testosterone and estradiol

No study provided data on measured free testosterone or estradiol. The gold standard method to measure free testosterone and free estradiol is via equiliburm dialysis [345, 346]. However, direct measurement of free hormones is slow, expensive and requires a relatively large sample volume [346]. Therefore, epidemiological

studies generally calculate free testosterone and estradiol. This analysis used the law of mass action to estimate free hormone concentrations [163, 302]. Although a number of other calculations exist [306, 347], this analysis chose to use the Sodergard calculation method, as this has been extensively validated using direct laboratory measurements and is most widely quoted in the literature [348–351]. Some studies have suggested that this calculation may perform less well in the presence of supraphysiologic concentrations of estradiol or DHT, or if men are receiving progestin or testosterone treatment [350, 352], these instances are rare. Other estimations of free testosterone have used statistical models to predict free testosterone concentrations and validated this against a measured sample [306, 347]. These empirical calculations are likely to be specific to their study population and laboratory methodology, whilst the calculations based on the law of mass action may be more generalisable. As a further analysis, free testosterone and free estradiol were calculated using alternative calculation methods [306], but this made no appreciable difference to the associations observed.

4.4.10 Strengths and limitations

This analysis has several strengths. The large size of the dataset enabled the assessment of associations across a wider range of exposure distributions than previously possible, particularly for A-diol-G and DHEAS, which have been the focus of less research. This analysis was also able to account for differences between studies to enable comparisons of the magnitudes of the associations relative to the different sex hormones.

This analysis has a number of limitations. The cross-sectional nature of the analysis means that it was not possible to determine whether these associations are causal. Furthermore, younger age groups and non-whites are under-represented. Data regarding other potential confounders such as physical activity, dietary factors, and co-morbidities were not collected in this centralised dataset, although only weak associations of these factors with hormone concentrations have been reported elsewhere [353, 354], therefore any related confounding of the studied associations

may be small. Analyses were based on a single measure of hormone concentrations, on the assumption that these measures are representative of participants' hormone concentrations over the medium-to-long term. While several studies show that single measures of these sex hormones have moderately good reproducibility over periods of up to five years [355–358] (e.g. for testosterone r=0.85 over a one year period) [355], the extent to which these measures are reproducible over the longer term is unclear.

4.5 Conclusion

In summary, this analysis of individual participant data from 25 studies shows age and BMI, and to a lesser extent, height, smoking status and alcohol consumption are associated with a range of circulating sex hormone concentrations. These associations will enable a greater understanding of how risk factors may relate to the development of hormone-related diseases.

5

Cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating IGF and IGFBP concentrations in men

Contents

5.1 Introduction 111
5.1.1 Aims 111
5.2 Subjects and methods 112
5.2.1 Data collection 112
5.2.2 Data processing 118
5.2.3 Statistical analysis 118
5.2.4 Multiple testing 120
5.3 Results 120
5.3.1 Correlations between IGF concentrations 120
5.3.2 Age at blood collection 122
5.3.3 Adiposity 124
5.3.4 Height 124
5.3.5 Smoking 129
5.3.6 Alcohol 129
5.3.7 Ethnic/racial group 129
5.3.8 Sociodemographic factors and health 133
5.3.9 Tests for heterogeneity and sensitivity analysis 133
5.3.10 Further analyses of IGFs and sex hormones 139
5.4 Discussion 141
5.4.1 Correlations between hormones 141
5.4.2 Age 141

5.4.3 Adiposity 142
5.4.4 Height 142
5.4.5 Smoking 143
5.4.6 Alcohol 143
5.4.7 Racial/ethnic group 143
5.4.8 Sociodemographic factors and health 144
5.4.9 Heterogeneity by study 144
5.4.10 Correlations between IGFs and sex hormones 144
5.4.11 Strengths and limitations 145
5.5 Conclusion 146

5.1 Introduction

Insulin-like growth factors (IGFs) and their associated binding proteins (IGFBPs) are fundamental mediators of cell growth, development and survival and are expressed in most tissues [210, 359] (Chapter 2). Epidemiological evidence indicates that the IGF system, particularly IGF-I, may be involved in prostate cancer risk [14], as well as breast, colorectal cancer [212, 213], and possibly thyroid cancer [214].

Risk of prostate cancer is associated with several anthropometric, behavioural, and sociodemographic factors, but the mechanisms conferring these associations are unclear (Chapter 2). Circulating concentrations of IGFs and IGFBPs may be modified by a range of anthropometric, behavioural, and sociodemographic factors [360, 361], many of which are also implicated in cancer aetiology [92, 362–365]. Therefore, IGFs provide a plausible biological mechanism through which these factors may be associated with cancer development. Although some previous research has investigated potential determinants of circulating IGF-I and IGFBP-3 [366–368], relatively few studies have investigated the correlates of IGF-II, IGFBP-1 and IGFBP-2 [366, 369].

5.1.1 Aims

Using the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG) dataset (Chapter 3) this analysis aims to examine the potential determinants of IGFs and IGFBPs. The application of

unified methods to this large individual-level participant dataset makes it possible to describe, with high precision, some of the potential determinants of IGFs and IGFBPs, both at extremes of their associations, and with substantial statistical power to identify novel associations. Identifying the correlates of circulating IGF and IGFBP concentrations will provide a better understanding of whether IGFs provide a plausible biological mechanism through which these factors may be associated with prostate cancer risk. This chapter also examines the partial correlations of circulating IGFs with sex hormone concentrations.

5.2 Subjects and methods

5.2.1 Data collection

The EHNBPCCG search strategy and pooling is described in greater detail in Chapter 3. Collaborators provided data on baseline IGF (IGF-I, IGF-II) and IGFBP (IGFBP-1, IGFBP-2, and IGFBP-3) concentrations. Assay details are available in Tables 5.1-5.2. Data on a range of participant baseline anthropometric, sociodemographic and behavioural factors were also provided. The data from each study were collected and incorporated into a central database.

Table 5.1: Assay methods and geometric mean IGF concentrations.

		IGF-I (nmol/L)			IGF-II (nmol/L)		
Study	Sample	Method	Geometric mean (95% CI)	CV %	Method	Geometric mean (95% CI)	CV %
ATBC	Serum	ELISA	17.7 (17.0-18.4)	6.6‡			
BLSA	Serum	E RIA	17.8 (16.6-19.0)	4.6-20‡	E RIA	40.8 (38.7-43.1)	4.9-30‡
BUPA	Serum	ELISA	15.0 (14.5-15.6)	-	ELISA	89.7 (87.3-92.2)	-
CHDS			30.2 (29.2-31.2)				
CHS	EDTA plasma	IRMA	19.2 (18.2-20.3)	3.0-12.3‡			
EPIC phase I	Serum	ELISA (excluding Swedish samples)	21.2 (20.6-21.8)	3.0-13.7‡			
EPIC phase III/IV** b	Serum	ELISA (IDS-iSYS Swedish samples)	19.1 (18.7-19.5)	3.2-4.4‡	Not published	111.6 (109.4-113.8)	
EPIC phase III/IV** b	Serum				Not published	59.3 (58.6-60.9)	
HIMS	Plasma	ELISA	16.8 (16.5-17.2)	8.6-12.2[a]			
HPFS Ib	Plasma	ELISA	22.2 (21.6-22.8)	<10 (batch 1998 to 2000, CV=13)†			
HPFS IIb	Plasma	ELISA	27.2 (26.4-27.9)				
JACC	Serum	IRMA	13.0 (12.1-14.0)	2.1-3.5†	IRMA	68.5 (64.9-72.8)	2.7-4.4†
KPMCP	Serum	E RIA	20.7 (19.7-21.7)	-			
MCCS	Plasma	ELISA	21.8 (21.4-22.3)	11.1[a]			
MEC	Serum	ELISA	22.9 (22.4-23.5)	2.1†	ELISA	114.4 (112.0-117.0)	1.8†
NSHDC	Plasma	E IRMA	25.3 (24.6-26.1)	8.6-13.8‡			
PCPT	Serum	ELISA	26.1 (25.5-26.7)	5.3-7.1†	ELISA	223.8 (219.9-227.7)	4.2-5.0†
PHS	Plasma	ELISA	22.7 (22.1-23.3)	4.9-6.5†	ELISA	66.5 (63.5-69.7)	-
PLCO	Serum	ELISA	24.9 (24.3-25.5)	9‡			
ProtecT feas*	Serum	ELISA	16.1 (15.7-16.6)	3-15‡	ELISA	54.1 (52.9-55.4)	5-26‡
ProtecT main	Serum	RIA	20.5 (20.2-20.9)	ICC 0.66-0.86	RIA	96.9 (95.6-98.2)	ICC 0.84- 0.91
SU.VI.MAX I	Plasma	CLIA	18.8 (18.1-19.5)	5.3¶	IRMA	142.4 (138.4-146.5)	6.8 ¶
SU.VI.MAX II **b			19.2 (18.4-19.9)			138.6 (134.4-143.0)	

* ProtecT feasibility study.
†Intra-assay.
a Inter-assay.
‡Intra-and inter-assay range.
¶ Not specified.
b Separate study assay populations.
* Not published.
Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; BUPA=British United Provident Association Study; CHDS=Child Health and Development Studies; CHS=Cardiovascular Health Study; CLIA=chemiluminescent immunometric assay; CV=coefficient of variation; E=extraction step; ECIA=electrochemiluminescence immunoassay; EDTA=ethylenediaminetetraacetic acid; ELISA=enzyme-linked immunosorbent assay; EPIC=European Prospective Investigation into Cancer and Nutrition; HIMS=Health In Men Study; HPFS=Health Professionals Follow-Up Study; ICC=intraclass correlations; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor; IRMA=Immunoradiometric assay; JACC=Japan Collaborative Cohort Study for Evaluation of Cancer; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=The Prostate, Lung, Colorectal and Ovarian; ProtecT=Prostate Testing for Cancer and Treatment; RIA=radioimmunoassay; SU.VI.MAX=SUpplémentation en VItamines et Minéraux AntioXydants.

Table 5.2: Assay methods and geometric mean IGFBP concentrations.

		IGFBP-1 (nmol/L)			IGFBP-2 (nmol/L)			IGFBP-3 (nmol/L)		
Study	Sample	Method	Geometric mean (95% CI)	CV %	Method	Geometric mean (95% CI)	CV %	Method	Geometric mean 95% CI)	CV %
BLSA	Serum							ELISA	79.0 (76.7-81.3)	7.30†
BUPA	Serum							NE RIA	94.8 (90.1-99.6)	5.1-17‡
CHDS	Serum							ELISA	95.2 (92.8-97.6)	-
CHS										
EPIC phase I	EDTA plasma							IRMA	108.2 (104-112.6)	2.1-7.1‡
EPIC phase III/IV** b	Serum	Not published	0.24 (0.22-0.26)					ELISA	128.8 (126.2-131.5)	5.30-9.40‡
EPIC phase III/IV** b	Serum				Not published	11.4 (10.9-11.9)				
HIMS		Not published	0.05 (0.05-0.05)		Not published	2.81 (2.73-2.90)				
HPFS Ib	Plasma	ELISA	0.78 (0.74-0.83)	5.2-8.6[a]				ELISA	128.7 (126.8-130.6)	4.4-16.8[a]
HPFS IIb									112.1 (109.8-114.3)	
JACC	Plasma	Not published	0.65 (0.59-0.71)					ELISA	126.7 (124.1-129.4)	<10†
KPMCP	Serum							IRMA	90.6 (85.8-95.6)	3.1-4.2†
MCCS	Serum									
MEC	Plasma							ELISA	103.6 (101.9-105.2)	9.5¶
NSHDC	Serum	ELISA	0.74 (0.68-0.80)	2.2†				ELISA	132.1 (129.7-134.6)	2.50†
PCPT	Plasma	IRMA	1.30 (1.44-1.48)	2.9†	RIA	16.4 (15.3-17.7)	2.50†	IRMA	80.2 (78.4-81.9)	3.6-6.9‡
PHS	Serum				ELISA	14.0 (13.5-14.5)	5.5-8.9‡	ELISA	136.6 (134.4-138.9)	4.2-4.8†
PLCO	Plasma	Not published	0.16 (0.15-0.18)					ELISA	107.8 (105.8-109.8)	7-9†
ProtecT feas*	Serum							ELISA	155.2 (152.4-158.0)	9†
ProtecT main	Serum				RIA	16.5 (15.6-17.3)	5-14‡	RIA	107.4 (105.0-109.8)	4-14‡
SU.VI.MAX I	Serum				ELISA	18.4 (17.9-18.9)	ICC 0.81-0.95‡	ELISA	149.5 (147.6-151.4)	ICC 0.71-0.88
SU.VI.MAX II **b	Plasma				RIA	6.6 (6.2-7.0)	8.6%¶	CLIA	143.0 (139.2-146.8)	6.3¶

*ProtecT feasibility study.
†Intra-assay.
a Inter-assay.
‡Intra-and inter-assay range.
¶ Not specified.
b Separate study assay populations.
* Not published.
Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; BUPA=British United Provident Association Study; CHDS=Child Health and Development Studies; CHS=Cardiovascular Health Study; CLIA=chemiluminescent immunometric assay; CV=coefficient of variation; E=extraction step; ECIA=electrochemiluminescence immunoassay; EDTA=ethylenediaminetetraacetic acid; ELISA=enzyme-linked immunosorbent assay; EPIC=European Prospective Investigation into Cancer and Nutrition; HIMS=Health In Men Study; HPFS=Health Professionals Follow-Up Study; ICC=intraclass correlations; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor; IRMA=immunoradiometric assay; JACC=Japan Collaborative Cohort Study for Evaluation of Cancer; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=The Prostate, Lung, Colorectal and Ovarian; ProtecT=Prostate Testing for Cancer and Treatment; RIA=radioimmunoassay; SU.VI.MAX=SUpplémentation en VItamines et Minéraux AntioXydants.

Inclusion criteria

Men were considered eligible for this analysis if they had measures of at least one of circulating IGF-I, IGF-II, IGFBP-1, IGFBP-2 or IGFBP-3 concentrations, were not known to have been diagnosed with prostate cancer by the time of individual study closure, and age, height and weight were recorded at the time of blood collection. Overall, there were 16,024 eligible men (out of 17,838; Figure 5.1) from the following studies: Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC) [229, 370], Baltimore Longitudinal Study of Aging (BLSA) [230, 231, 371], British United Provident Association Study (BUPA) [232], Child Health and Development Studies (CHDS) [234, 372], Cardiovascular Health Study (CHS) [235], European Prospective Investigation into Cancer and Nutrition (EPIC) [237, 238], Health in Men Study (HIMS) [241, 242, 373, 374], Health Professionals Follow-up Study (HPFS) [375, 376], Japan Collaborative Cohort Study (JACC) [18, 246], Kaiser Permanente Medical Care Program (KPMCP) [251], Melbourne Collaborative Cohort Study (MCCS) [253], Multiethnic Cohort (MEC) [254], Northern Sweden Health and Disease Cohort (NSHDC) [244, 377], Prostate Cancer Prevention Trail (PCPT) [10, 378], Physicians' Health Survey (PHS) [256, 379–381], Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) [382], Prostate Testing for Cancer and Treatment (ProtecT) feasibility study [258, 259] and main study [260], SUpplémentation en VItamines et Minéraux AntioXydants (SU. VI. Max) [262], by dataset closure on 1st September, 2018.

The characteristics of men included in this analysis are summarised in Table 5.3.

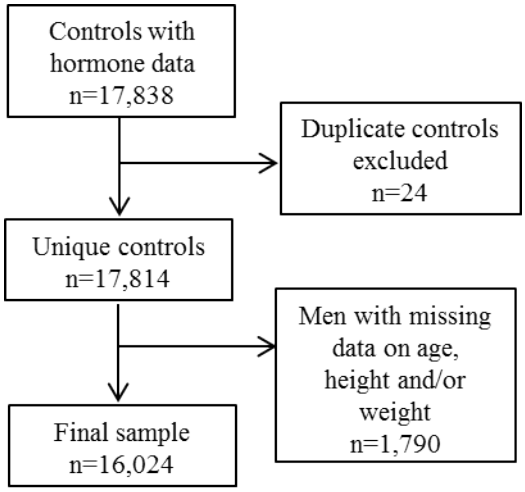


Figure 5.1: Participant selection chart.

Table 5.3: Participant characteristics by studies with IGF data

Study, country	N (% of total)	Mean age, years (SD)	Age range, years	Year of birth	Mean height, cm (SD)	Mean BMI, kg/m^2 (SD)	% Current drinkers (median daily alcohol consumption, g alcohol)	% Current smokers (median number of daily cigarettes)	% White ethnic group	% Married/ cohabiting at blood collection	% University degree	% Family history of prostate cancer
ATBC, Finland	311 (1.9)	58.5 (4.5)	52-70	1916-1935	173.5 (5.9)	26.5 (3.9)	82.2 (10)	100 (20)	100	81.4	5.5	4.2
BLSA, USA	110 (0.7)	64.7 (9.4)	43-83	1895-1944	175.0 (6.7)	26.5 (3.7)	-	7.3	97.3	83.3	58.2	-
BUPA, UK	423 (2.6)	54.5 (6.2)	36-64	1911-1941	175.6 (7.0)	25.4 (2.9)	98.1 (15)	19.1 (20)	-	-	-	-
CHDS, USA	434 (2.7)	33.9 (6.9)	22-50	1912-1943	178.6 (6.7)	24.6 (2.7)	76.8 (5)	52.0 (20)	62.4	99.8	33.1	-
CHS, USA	174 (1.1)	72.4 (4.4)	65-89	1902-1926	173.5 (6.6)	26.7 (4.1)	-	13.8	75.7	83.3	14.5	-
EPIC phase I, Europe	636 (4.0)	60.9 (6.2)	43-76	1918-1952	172.6 (7.0)	27.0 (3.6)	87.2 (13)	27.9 (15)	100	89.3	23	-
EPIC phase II, Europe	1,193 (7.4)	58.7 (6.1)	39-77	1918-1956	174.3 (6.8)	26.7 (3.6)	90.5 (13)	25.9 (15)	100	88.5	24.1	-
EPIC phase III/IV‡, Europe	1,787 (11.2)	56.1 (7.2)	36-78	1918-1958	172.5 (7.2)	26.9 (3.4)	88.8 (14)	25.3 (16)	100	88.4	23.1	-
HIMS, Australia	1,279 (8.0)	76.3 (3.6)	71-87	1915-1932	171.7 (6.9)	26.5 (3.7)	65.9 (7)	4.8	100	86.3	21.4	-
HPFS I, USA	682 (4.3)	65.1 (7.4)	46-80	1913-1946	178.1 (6.4)	26.0 (3.5)	73.5 (6)	3.9 (20)	99.4	93	100	10.3
HPFS II, USA	629 (3.9)	62.0 (7.8)	46-80	1914-1946	177.5 (6.7)	26.1 (3.6)	71.5 (6)	3	92.4	93	100	10.8
JACC, Japan	94 (0.6)	68.1 (5.6)	58-83	1904-1931	159.4 (6.9)	22.4 (2.7)	52.1 (2)	37.0 (20)	0	93.4	2.7	-
KPMCP, USA	212 (1.3)	71.8 (4.5)	60-85	1882-1906	169.9 (6.7)	25.8 (3.1)	69.7 (10)	17.9 (30)	98.6	82.7	5.4	-
MCCS, Australia	1,047 (6.5)	58.3 (7.2)	40-72	1918-1954	172.1 (7.3)	27.2 (3.7)	78.3 (13)	13.2 (20)	100	81	22.2	-
MEC, USA	772 (4.8)	68.5 (7.1)	49-84	1918-1948	173.9 (7.7)	26.9 (4.1)	90.5 (9)	11.7 (15)	14.5	79.2	31.9	8.3
NSHDC, Sweden	557 (3.5)	57.9 (4.3)	40-72	1921-1948	175.4 (5.9)	26.6 (3.7)	100 (4)	20.8	100	80	12.3	-
PCPT, USA and Canada	1,022 (6.4)	63.3 (5.5)	55-83	1911-1940	177.5 (7.0)	27.6 (4.0)	69.0 (3)	7.6 (20)	84	87.7	37	20.7
PHS, USA	757 (4.7)	58.4 (8.0)	39-84	1898-1942	178.1 (6.8)	24.7 (2.5)	83.6 (5)	8.9 (20)	94.2	-	100	15.4
PLCO, USA	858 (5.4)	64.8 (4.8)	54-75	1919-1944	177.9 (6.5)	27.4 (3.9)	70.0 (4)	9.1 (20)	100	86.9	42.4	6.1
ProtecT feas*, UK	568 (3.5)	61.5 (5.0)	50-70	1930-1949	175.3 (6.5)	26.6 (3.6)	83.5 (17)	10.9	-	-	-	4.1
ProtecT main, UK	1,770 (11.0)	61.9 (5.0)	45-70	1933-1958	176.1 (6.5)	26.9 (3.7)	84.6 (17)	13.6	-	-	-	5.6
SU.VI.MAX I, France	382 (2.4)	55.0 (4.6)	44-61	1933-1950	174.1 (6.1)	25.4 (2.9)	84.2 (28)	12.9	-	87.8	34.8	5.1
SU.VI.MAX II‡, France	327 (2.0)	53.3 (4.3)	42-61	1932-1952	172.6 (6.3)	25.7 (3.2)	80.7 (28)	13.1 (18)	100	89.8	26.5	4.3
Overall†	16,024	61.2 (9.5)	22-89	1882-1958	174.7 (7.3)	26.6 (3.7)	80.6 (10)	17.0 (20)	90.4	87	38.1	8.8
Missing data	-	-	-	-	-	-	7.8	1.3	19.9	31.4	19.7	56.8

*ProtecT feasibility study.
† Excludes missing data.
a Excludes studies where data were not collected.
‡ Not published.
Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BMI=body mass index; BLSA=The Baltimore Longitudinal Study of Aging; BUPA=British United Provident Association Study; CHDS=Child Health and Development Studies; CHS=Cardiovascular Health Study; EPIC=European Prospective Investigation into Cancer and Nutrition; HIMS=Health In Men Study; HPFS=Health Professionals Follow-up Study; IGF=insulin-like growth factor; JACC=Japan Collaborative Cohort Study for Evaluation of Cancer; KPMCP=Kaiser Permanente Medical Care Program; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=The Prostate, Lung, Colorectal and Ovarian; ProtecT=Prostate Testing for Cancer and Treatment; SD=standard deviation; SU. VI MAX=SUpplémentation en VItamines et Minéraux AntioXydants; UK=United Kingdom; USA=United States of America.

5.2.2 Data processing

IGF-I, IGF-II, IGFBP-1, IGFBP-2 and IGFBP-3 concentrations were log-transformed to approximate normal distributions. Factors investigated were age, body mass index (BMI), waist circumference, waist-to-hip ratio (WHR), height, smoking status, alcohol consumption, ethnic/racial group, educational status, marital status, and family history of prostate cancer. Variable categories were as defined in Chapter 4 (Section 4.2.2).

5.2.3 Statistical analysis

Statistical methodologies are described in greater detail in Chapter 4.

In brief, partial correlations between the IGFs and IGFBPs were calculated using study-specific standardised values and adjusted for age at blood collection, BMI and height (included as categorical variables, categorised in Chapter 4).

Geometric mean concentrations of IGFs and IGFBPs were calculated, and adjusted for study, age at blood collection, BMI and height. In these analyses height was included as an adjustment covariate due to the well-established role of IGF-I in regulating bone growth [94, 383]. When the associations of age, height and BMI were analysed with hormone concentrations, these variables were not included as adjustment covariates. Adjusted geometric mean concentrations for waist circumference and WHR were also repeated with and without adjustment for BMI. Analyses of smoking and alcohol consumption were mutually adjusted for each other. To investigate the relationship of IGF and IGFBP concentrations with ethnicity/race, the analyses were limited to the five (all USA-based) studies that had sufficient representation from men across several ethnic/racial groups (CHDS, CHS, MEC, PCPT and PHS).

To enable robust adjustment for study, each study had to contain observations in a minimum of two categories for each primary exposure to be included in the respective exposure analysis.

Heterogeneity of means by category of each characteristic was tested using the F test. The test for trend was calculated using the analysis of variance test, with

the categorical variables entered as linear values scored consecutively as 1, 2, 3 etc. For alcohol consumption, the test for trend was calculated based on median values within each category, excluding non-drinkers. To test for trend by smoking status, never and former smokers were combined and coded as 0; light, medium and heavy smokers were coded as 1, 2 and 3, respectively. In a secondary analysis, the test for trend was calculated for current smokers only.

Heterogeneity between studies was tested using a study-by-factor interaction term (fitted separately) in the analysis of variance and assessed using the F test. For the analysis of BMI, heterogeneity in the associations between the studies which measured and used self-reported BMI was assessed using the likelihood ratio test. Circulating IGFBP-1 and IGFBP-2 concentrations are known to be affected by fasting status [384, 385]. As fasting status was not recorded for 58% of participants, this variable was not included as a covariate in the analyses, but heterogeneity between exposure factors and overnight fasting status for these two binding proteins was assessed using the likelihood ratio test.

Sensitivity analysis

Sensitivity analyses were conducted after restricting the dataset to: i) white men only (n=11,611), ii) studies which used enzyme-linked immunosorbent assays (ELISA), iii) men with IGF and IGFBP concentrations that were within the range of [lower quartile – 3*interquartile range, upper quartile + 3*interquartile range] within each study in order to examine the effect of outliers (n=147). IGFs are also determinants of bone growth during development, therefore it is possible that adjusting for height in this analysis may attenuate the associations, therefore the primary analysis was repeated without adjustment for height. The primary analysis was also repeated after further adjustment for smoking and alcohol.

Further analyses of IGFs and sex hormones

Twelve studies contributed measurements of both circulating sex hormones and IGFs to the EHNBPCCG (CHDS, EPIC, HIMS, HPFS, JACC, KPMCP, MCCS, MEC, PCPT, PHS, PLCO and ProtecT feasibility; up to 7,759 men). As a further

analysis, partial correlations between log-transformed sex hormones and IGFs were calculated by standardising the measures by study and adjusting for age at blood collection, BMI and height.

5.2.4 Multiple testing

This analysis contains a large number of statistical tests, which increases the probability of type I errors. Because the analyses were exploratory, the analyses did not adjust further for multiple testing to avoid rejecting potentially important findings [307]. However, the associations observed, particularly those with weaker p-values, have been reported in the context of evidence from other studies.

Statistical software

All statistical tests were two-sided. Data analyses used Stata Statistical Software release 14.1 (Stata Corp., College Station, TX, USA), and figures were created in R version 3.2.3.

5.3 Results

Twenty-three studies including 16,024 participants contributed data to the analysis (Table 5.3). Participants were born between 1882-1958. Blood was collected 1959-2009, and age at blood collection ranged from 22-89 years (mean=61.2; SD=9.5 years). Participants were predominantly non-Hispanic white (90%), non-smokers (83%), and 81% reported consuming some alcohol (median per day=10g ethanol, approximately equal to one standard alcoholic beverage).

5.3.1 Correlations between IGF concentrations

After standardising for study and adjusting for age, BMI and height, all IGFs and IGFBPs were correlated with each other. IGF-I was positively correlated with both IGFBP-3 and IGF-II (r=0.58 and 0.41, respectively) and weakly inversely correlated with IGFBP-1 and IGFBP-2 (r=-0.15 and -0.09, respectively). IGF-II

was positively correlated with IGFBP-3 (r=0.65) and inversely correlated with IGFBP-1 and IGFBP-2 (r=-0.20 and -0.11, respectively) (Figure 5.2).

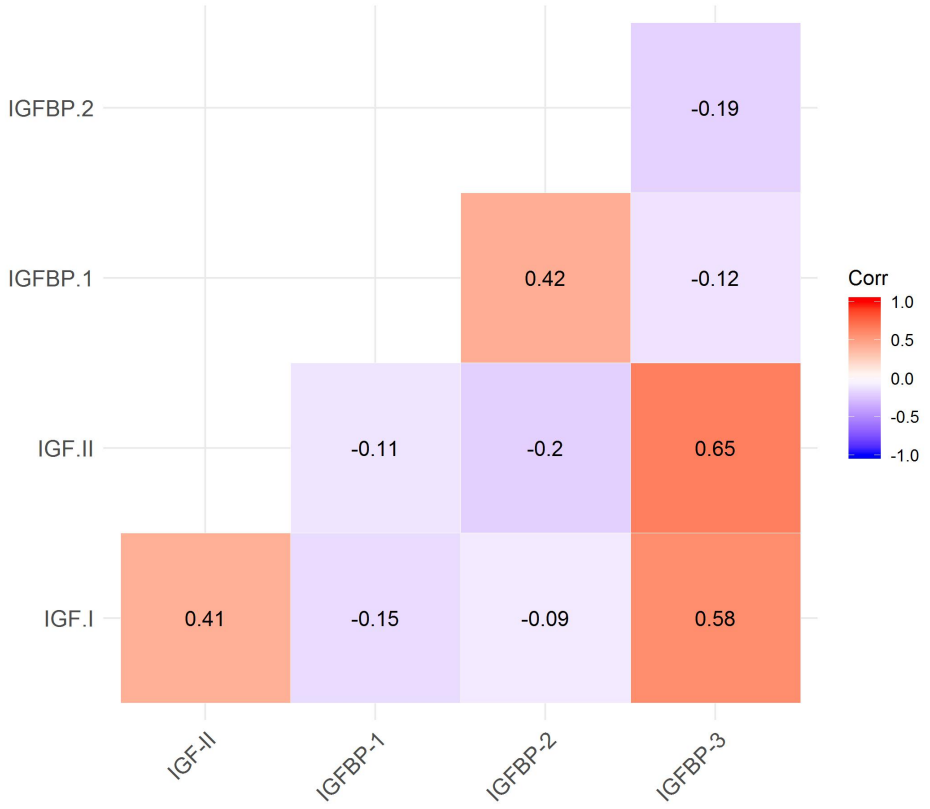


Figure 5.2: Partial correlation coefficients in control subjects between log-transformed IGFs and IGFBPs.
Values are standardised by study and adjusted by age, BMI and height.
Number of observations ranged from 12,012 (IGF-I and IGFBP-3) to 2,873 (IGFBP-1 and IGFBP-2).
Abbreviations: BMI=body mass index; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor-binding protein.

5.3.2 Age at blood collection

After adjusting for study, BMI and height, age was associated with all IGFs and IGFBPs (Figure 5.3). Compared with men aged 50-54 years, those aged 75+ years had circulating concentrations of IGFBP-2 and IGFBP-1 that were 73% and 61% higher, respectively, whilst IGF-I, IGFBP-3 and IGF-II concentrations were 24%, 21% and 19% lower, respectively.

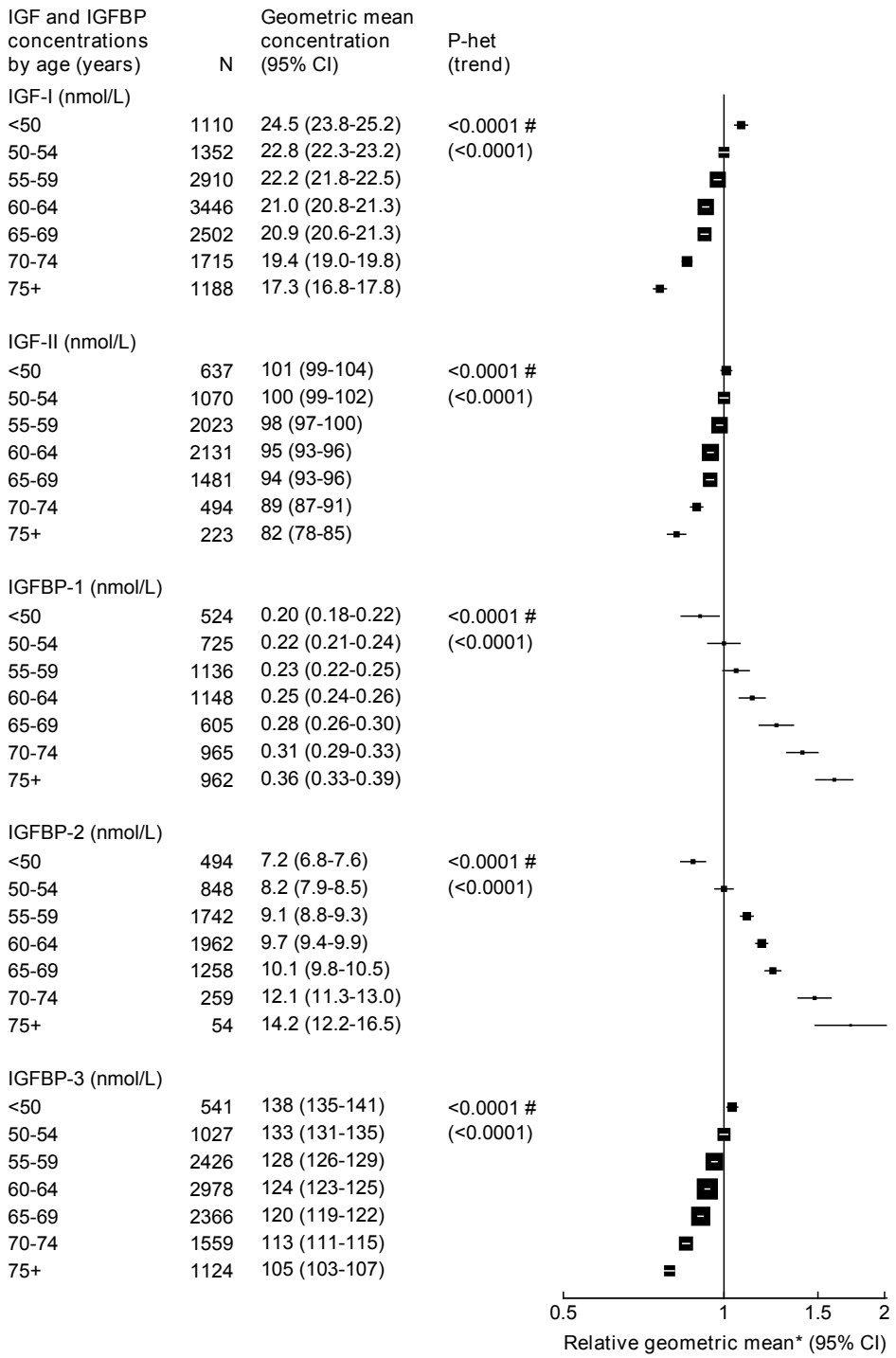


Figure 5.3: Relative geometric mean concentrations* of male IGFs and IGFBPs by age. P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
*relative to 50-54 years, adjusted for study, BMI and height.
#significant heterogeneity by study ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

5.3.3 Adiposity

After adjusting for study, age and height, BMI was associated with IGF and IGFBP concentrations (Figure 5.4). Compared with men with a BMI of 20.0-22.4 kg/m^2, those with a BMI 37.5+ kg/m^2 had concentrations of IGFBP-1 and IGFBP-2 that were 77% and 58% lower, respectively. IGF-I, IGF-II and IGFBP-3 had inverse U-shaped associations with BMI.

The patterns of associations were broadly similar for waist circumference and WHR (Figures 5.5 and 5.6), although there was no evidence of an association between WHR and IGF-I or IGFBP-3. Compared with men with waist circumference 90-94 cm, those with waist circumference 105+ cm had circulating concentrations of IGFBP-1 and IGFBP-2 that were 43% and 37% lower, respectively (Figure 5.5). Following additional adjustment for BMI, IGFBP-1 and IGFBP-2 concentrations were 19% and 14% lower, respectively in men with waist circumference 105+ cm. Compared with men with WHR 0.900-0.932, those with WHR 1.000+ had circulating concentrations of IGFBP-1 and IGFBP-2 that were 34% and 30% lower, respectively (Figure 5.6). After additional adjustment for BMI, IGFBP-1 and IGFBP-2 concentrations were 11% and 10% lower, respectively in men with WHR 1.000+.

There was evidence of heterogeneity in the associations with BMI by self-report and measured methodologies. This appeared to be due to differences in the magnitude rather than the direction of associations (data not shown).

5.3.4 Height

After adjusting for study, age and BMI, height was associated with concentrations of IGF-I, IGFBP-1 and IGFBP-3 (Figure 5.7). Compared with shorter men (160-164 cm), the tallest men (190+ cm) had concentrations of IGFBP-1 that were 35% lower, whilst IGF-I and IGFBP-3 concentrations were 7% and 3% higher, respectively.

IGF and IGFBP concentrations by BMI (kg/m²)	N	Geometric mean concentration (95% CI)	P-het (trend)
IGF-I (nmol/L)			
<20.0	197	19.2 (18.2-20.2)	<0.0001
20.0-22.4	1363	20.3 (19.9-20.7)	(0.0058)
22.5-24.9	3596	21.4 (21.2-21.7)	
25.0-27.4	4274	21.5 (21.2-21.7)	
27.5-29.9	2654	21.3 (21.0-21.6)	
30.0-32.4	1253	20.6 (20.2-21.0)	
32.5-34.9	521	20.0 (19.4-20.7)	
35.0-37.4	216	19.2 (18.3-20.1)	
37.5+	149	19.0 (17.9-20.1)	
IGF-II (nmol/L)			
<20.0	114	88 (84-93)	<0.0001 #
20.0-22.4	684	93 (91-95)	(0.0216)
22.5-24.9	1936	95 (93-96)	
25.0-27.4	2483	98 (97-99)	
27.5-29.9	1562	97 (96-99)	
30.0-32.4	763	97 (95-99)	
32.5-34.9	317	95 (93-98)	
35.0-37.4	115	92 (88-97)	
37.5+	85	89 (83-94)	
IGFBP-1 (nmol/L)			
<20.0	72	0.54 (0.44-0.67)	<0.0001 #
20.0-22.4	554	0.47 (0.44-0.51)	(<0.0001)
22.5-24.9	1521	0.34 (0.33-0.36)	
25.0-27.4	1871	0.26 (0.25-0.27)	
27.5-29.9	1179	0.20 (0.19-0.21)	
30.0-32.4	547	0.18 (0.16-0.19)	
32.5-34.9	184	0.16 (0.14-0.18)	
35.0-37.4	75	0.11 (0.09-0.13)	
37.5+	62	0.11 (0.09-0.14)	
IGFBP-2 (nmol/L)			
<20.0	74	18.2 (16.0-20.8)	<0.0001
20.0-22.4	509	15.4 (14.7-16.2)	(<0.0001)
22.5-24.9	1506	12.2 (18.9-12.6)	
25.0-27.4	2049	9.1 (8.9-9.3)	
27.5-29.9	1347	7.7 (7.4-7.9)	
30.0-32.4	673	6.6 (6.4-6.9)	
32.5-34.9	274	6.2 (5.8-6.7)	
35.0-37.4	107	5.8 (5.1-6.4)	
37.5+	78	6.5 (5.7-7.3)	
IGFBP-3 (nmol/L)			
<20.0	159	113 (108-117)	<0.0001
20.0-22.4	1107	117 (116-119)	(0.8213)
22.5-24.9	3018	123 (122-124)	
25.0-27.4	3583	124 (123-125)	
27.5-29.9	2280	122 (121-124)	
30.0-32.4	1081	120 (119-122)	
32.5-34.9	472	120 (117-123)	
35.0-37.4	189	118 (114-123)	
37.5+	132	115 (110-120)	

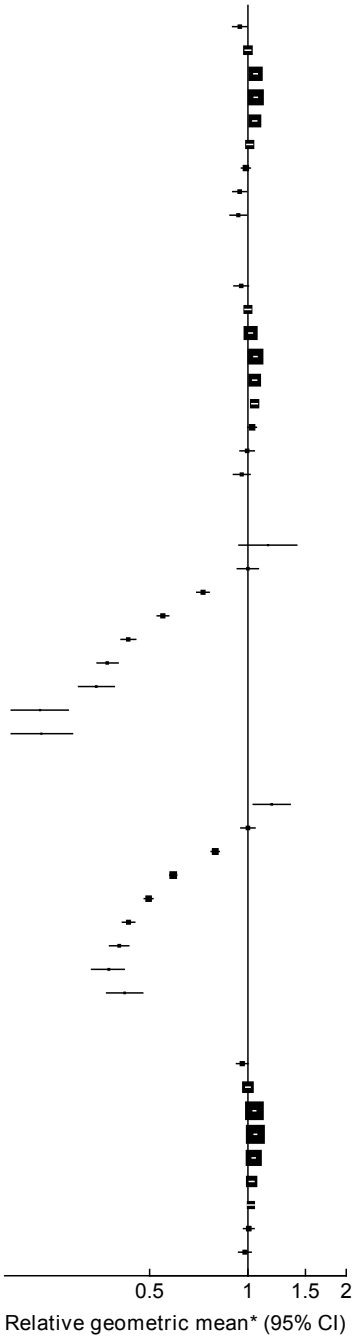


Figure 5.4: Relative geometric mean concentrations* of male IGFs and IGFBPs by BMI.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
*relative to BMI 20.0-22.4 kg/m^2, adjusted for study, age and height.
#significant heterogeneity by study ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

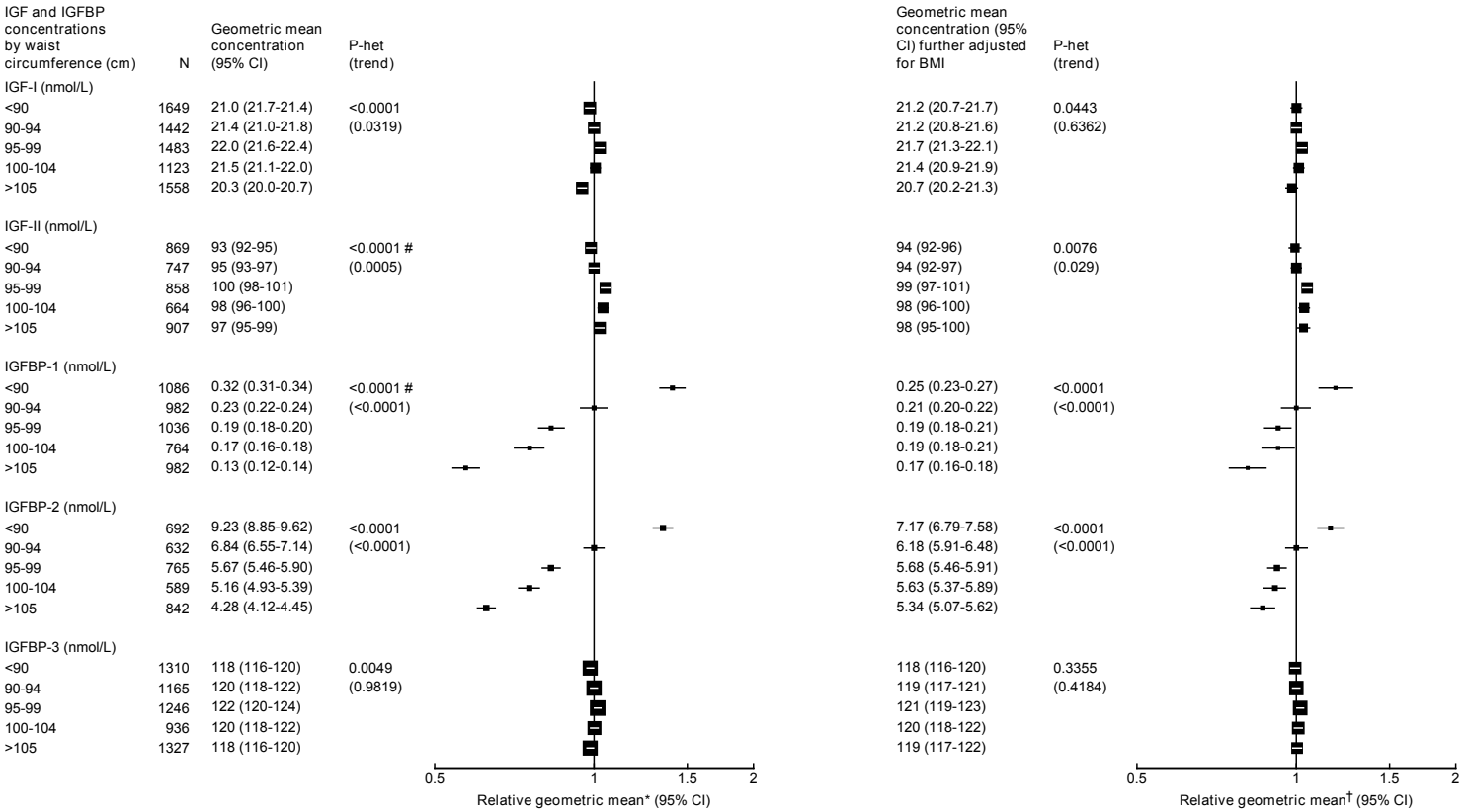


Figure 5.5: Relative geometric mean concentrations* of male IGFs and IGFBPs by waist circumference.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
*relative to 90-94 cm, adjusted for study, age and height
† relative to 90-94 cm, adjusted for study, age, height and BMI.
#significant heterogeneity by study ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

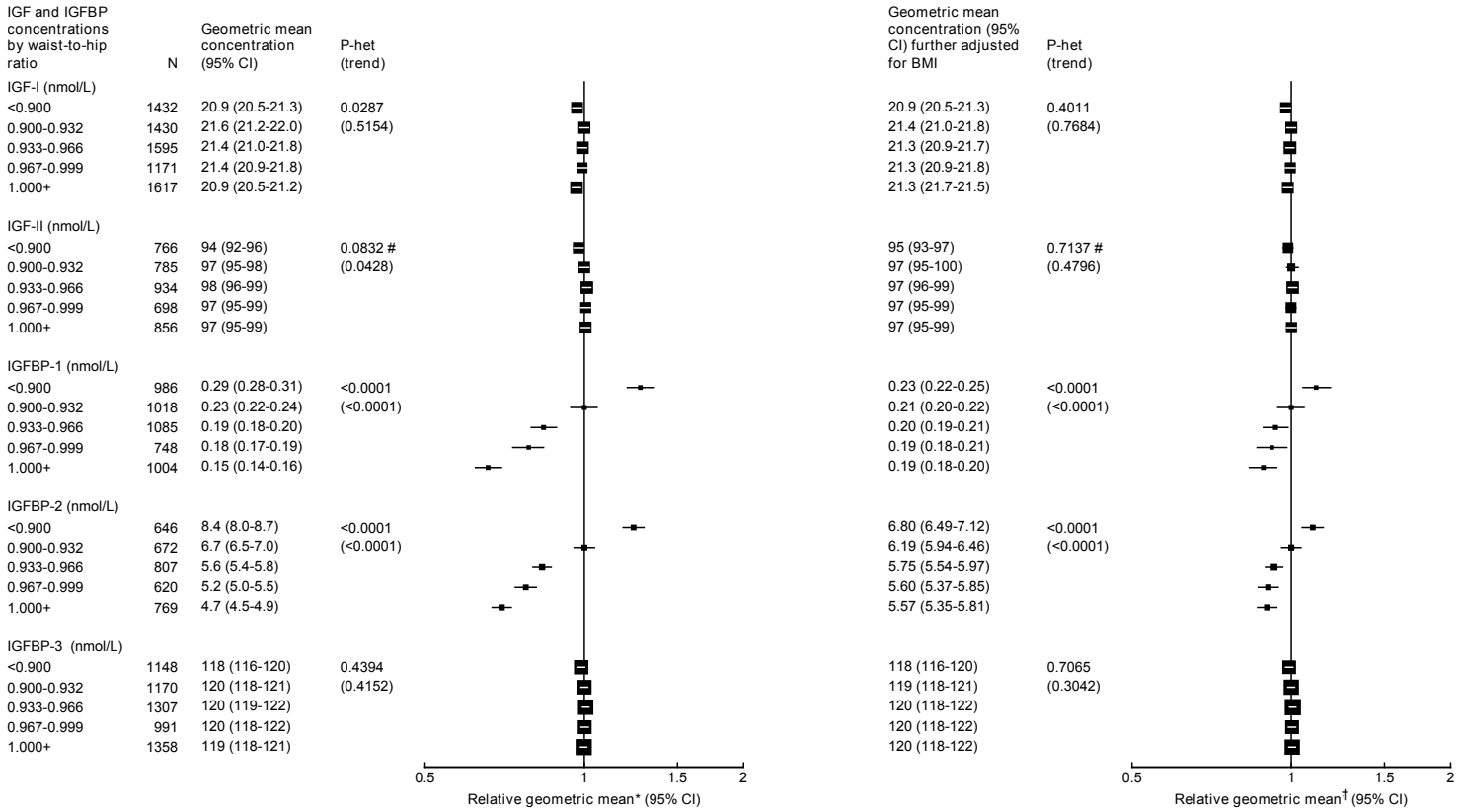


Figure 5.6: Relative geometric mean concentrations* of male IGFs and IGFBPs by WHR.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
*relative to 0.900-0.932, adjusted for study, age and height.
† relative to 0.900-0.932, adjusted for study, age, height and BMI.
#significant heterogeneity by study ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein; WHR=waist-to-hip-ratio.

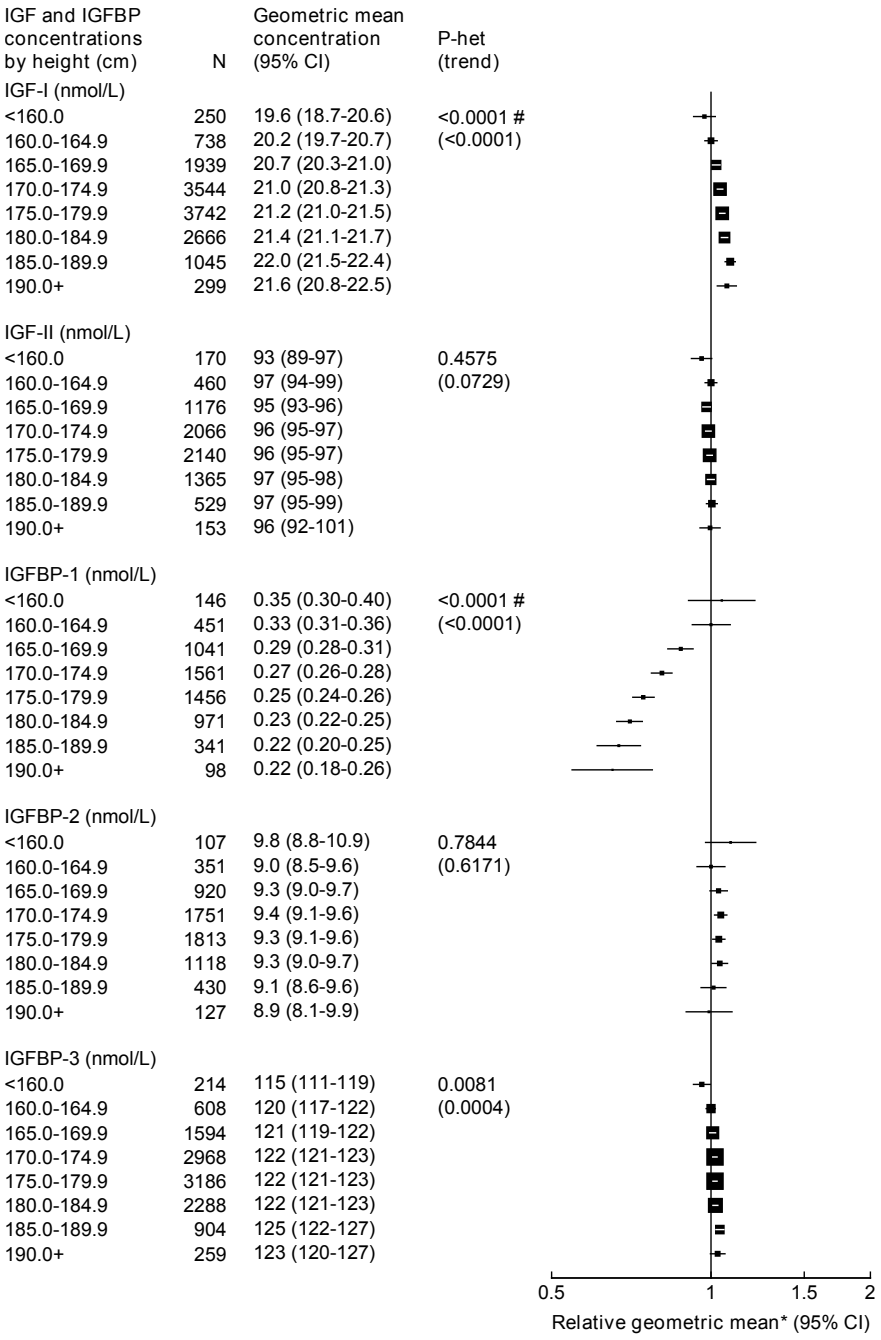


Figure 5.7: Relative geometric mean concentrations* of male IGFs and IGFBPs by height.

P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.

*relative to 175.0-179.9 cm, adjusted for study, age and BMI.

#significant heterogeneity by study ($p<0.01$).

Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

5.3.5 Smoking

After adjusting for study, age, BMI, height and alcohol consumption, smoking status was associated with IGFBP concentrations (Figure 5.8). Compared with never smokers, heavy smokers (30+ cigarettes/day) had concentrations of IGFBP-1 and IGFBP-2 that were 39% and 20% higher, respectively, whilst heavy smokers had 4% lower concentrations of IGFBP-3. However, there was no evidence of associations between smoking intensity and circulating IGFBPs when the analyses were restricted to current smokers only (data not shown).

5.3.6 Alcohol

After adjusting for study, age, BMI, height and smoking status, alcohol consumption was associated with IGFs and IGFBP-2 (Figure 5.9). Compared with men who drank 1-9 g ethanol/day, men who drank 70+ g ethanol /day had circulating concentrations of IGF-II that were 5% higher, whilst IGF-I and IGFBP-2 concentrations were 13% and 7% lower, respectively.

Non-drinkers had 5%, 4% and 2% lower concentrations of IGF-II, IGFBP-3 and IGF-I, respectively in comparison with men who drank 1-9 g ethanol per day.

5.3.7 Ethnic/racial group

After adjusting for study, age, BMI and height, racial/ethnic group was associated with all IGF and IGFBP concentrations (Figure 5.10). Compared with non-Hispanic white men, African American men had concentrations of IGFBP-1, IGFBP-2, IGF-II, and IGFBP-3 that were 36%, 33%, 11% and 6% lower, respectively. Hispanics had concentrations of IGFBP-2, IGF-I, IGFBP-3 and IGF-II that were 15%, 13%, 12% and 9% lower, respectively. Concentrations of IGFs and IGFBPs for East Asian men were similar to those for non-Hispanic white men.

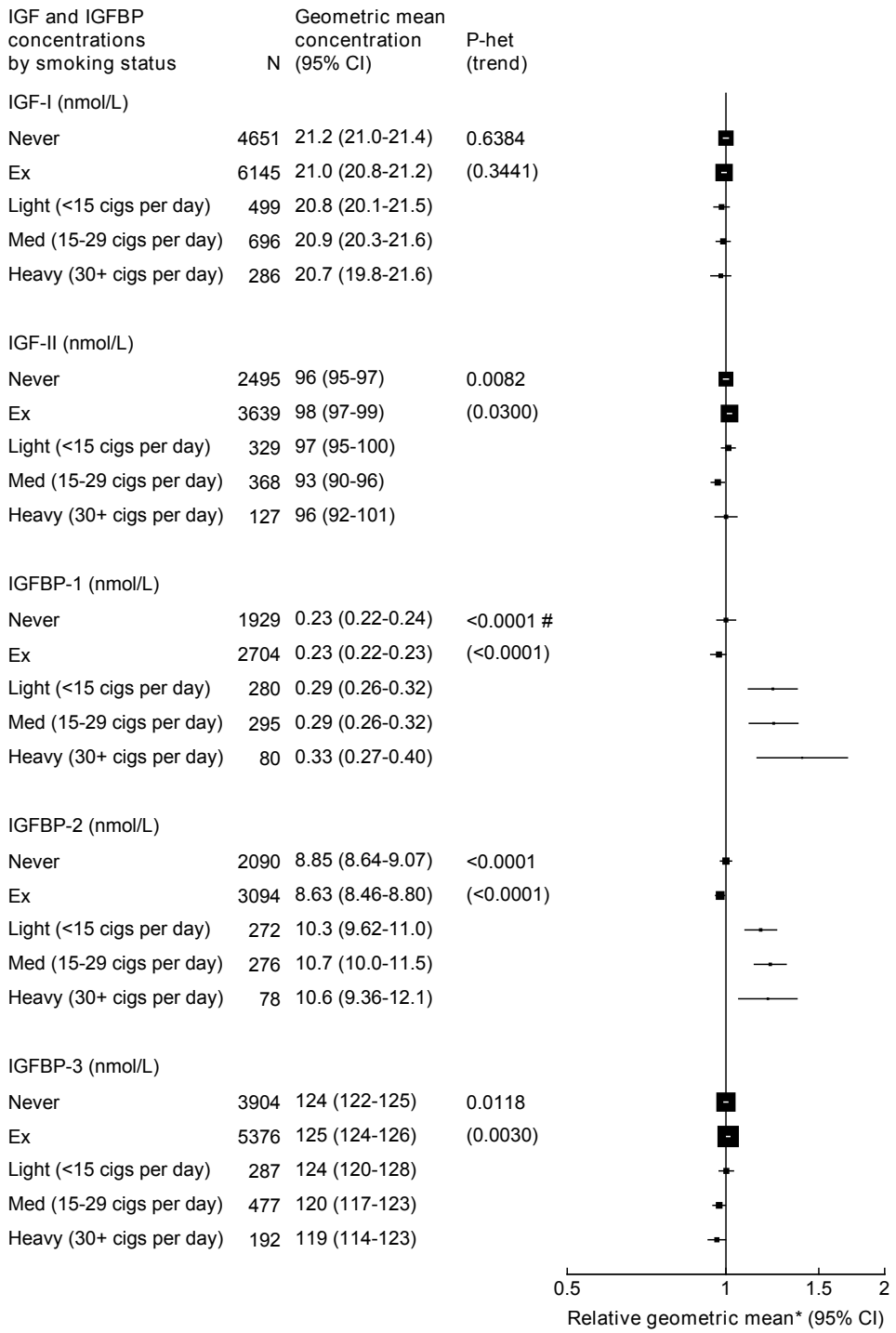


Figure 5.8: Relative geometric mean concentrations* of male IGFs and IGFBPs by smoking status.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively. Never and ex-smokers were combined into a single category.
*relative to never smokers, adjusted for study, age, BMI, height and alcohol consumption.
#significant heterogeneity by study ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

IGF and IGFBP concentrations by drinking status (g ethanol/day)	N	Geometric mean concentration (95% CI)	P-het (trend)
IGF-I (nmol/L)			
None	2558	21.2 (20.9-21.5)	<0.0001 #
1-9	3481	21.7 (21.4-21.9)	(<0.0001)
10-19	2439	21.4 (21.1-21.7)	
20-29	1381	21.0 (20.6-21.5)	
30-39	832	20.2 (19.7-20.7)	
40-49	633	20.5 (19.9-21.1)	
50-59	399	20.0 (19.3-20.7)	
60-69	226	19.5 (18.6-20.4)	
70+	327	19.0 (18.2-19.7)	
IGF-II (nmol/L)			
None	1085	92 (91-94)	<0.0001
1-9	1848	97 (96-98)	(<0.0001)
10-19	1252	98 (97-100)	
20-29	982	99 (97-100)	
30-39	617	99 (97-101)	
40-49	486	102 (100-105)	
50-59	277	105 (101-108)	
60-69	187	101 (97-105)	
70+	224	102 (99-106)	
IGFBP-1 (nmol/L)			
None	971	0.24 (0.23-0.26)	0.4882
1-9	1507	0.23 (0.22-0.24)	(0.2920)
10-19	1155	0.23 (0.22-0.24)	
20-29	625	0.24 (0.23-0.26)	
30-39	386	0.22 (0.20-0.24)	
40-49	253	0.22 (0.20-0.25)	
50-59	143	0.24 (0.21-0.28)	
60-69	117	0.26 (0.22-0.30)	
70+	131	0.25 (0.21-0.29)	
IGFBP-2 (nmol/L)			
None	930	9.08 (8.75-9.43)	0.0065
1-9	1522	9.00 (8.75-9.27)	(0.0009)
10-19	1037	8.84 (8.54-9.15)	
20-29	821	8.52 (8.20-8.87)	
30-39	554	8.26 (7.88-8.66)	
40-49	354	8.34 (7.86-8.85)	
50-59	243	8.19 (7.63-8.80)	
60-69	151	8.35 (7.63-9.15)	
70+	198	8.37 (7.73-9.06)	
IGFBP-3 (nmol/L)			
None	2238	119 (118-121)	<0.0001
1-9	2916	124 (123-126)	(0.2603)
10-19	1995	125 (124-127)	
20-29	1098	124 (122-126)	
30-39	688	125 (123-127)	
40-49	531	127 (124-130)	
50-59	317	126 (123-130)	
60-69	176	128 (124-133)	
70+	277	123 (120-127)	

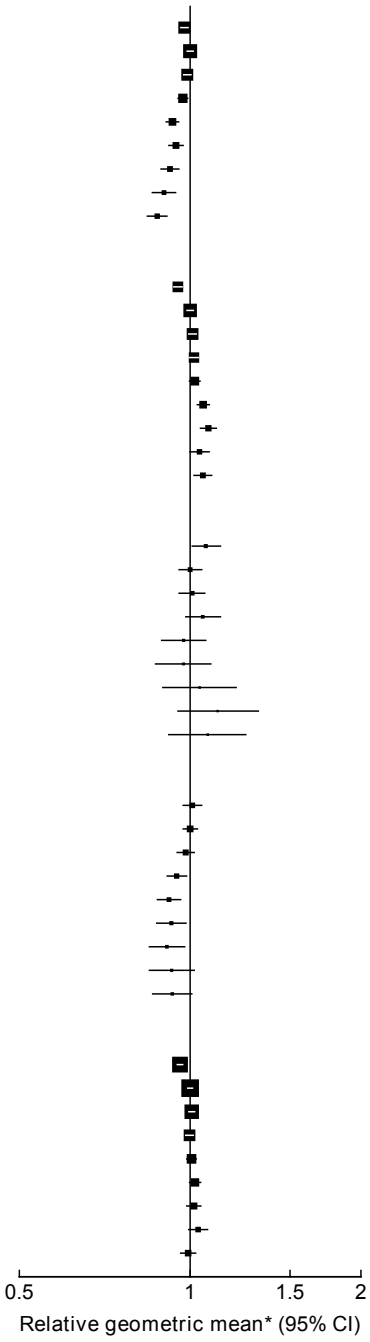


Figure 5.9: Relative geometric mean concentrations* of male IGFs and IGFBPs by alcohol consumption.

P for heterogeneity is the heterogeneity of means between categories, tested using the F test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored as the median value within each category and excluded non-drinkers.

*relative to non-drinkers, adjusted for study, age, height, BMI and smoking status.

#significant heterogeneity by study ($p<0.01$).

Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

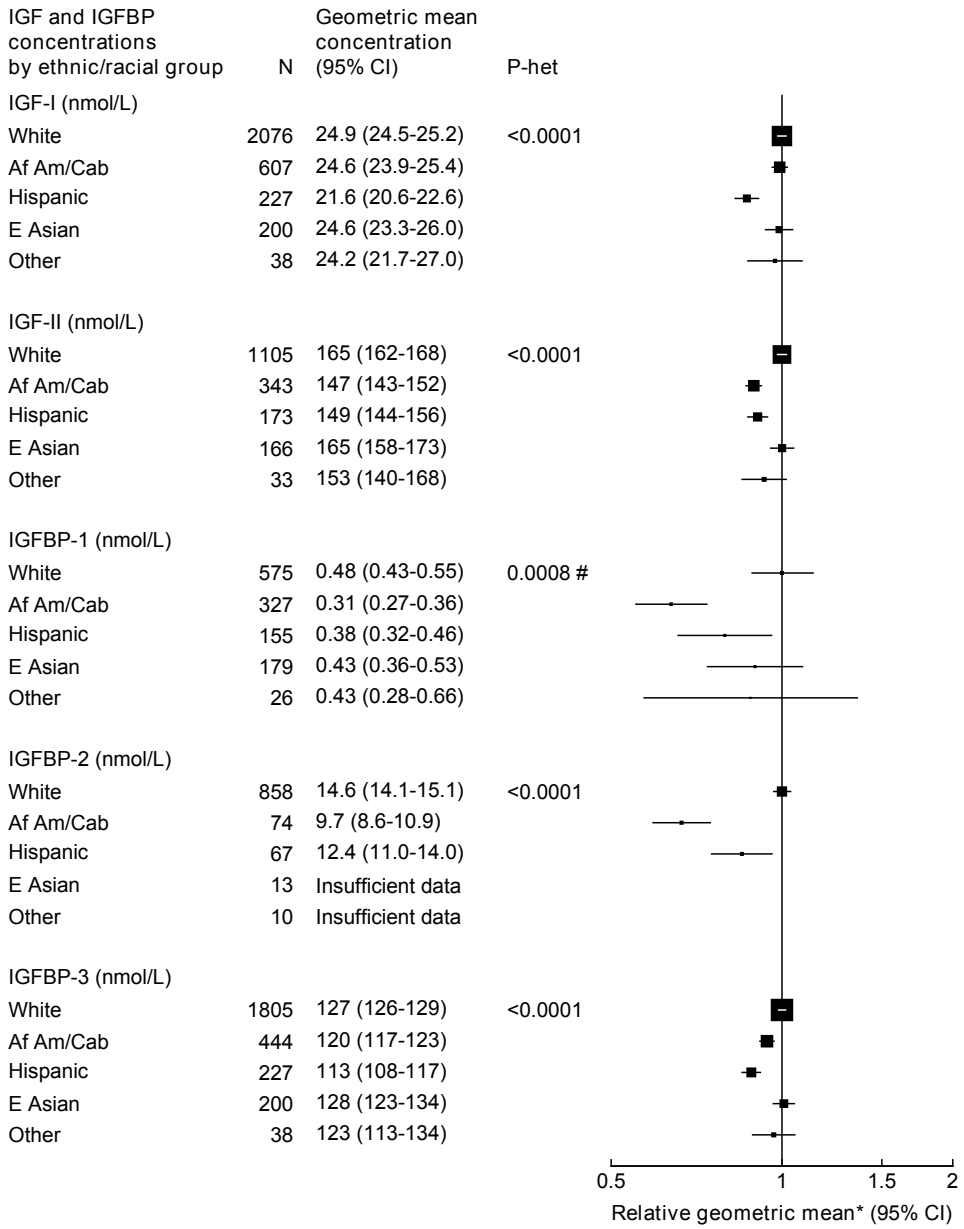


Figure 5.10: Relative geometric mean concentrations* of male IGFs and IGFBPs by racial/ethnic group.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test.
*relative to non-Hispanic whites, adjusted for study, age, height, and BMI.
#significant heterogeneity by study ($p<0.01$).
Abbreviations: Af Am/Cab=African American/Caribbean; BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

5.3.8 Sociodemographic factors and health

After adjusting for study, age, BMI and height, men who were married or cohabiting at the time of blood collection had 4% higher concentrations of IGF-I and IGF-II compared with men who were not married at blood collection (Figure 5.11). Men with a university education had 6%, 5% and 3% higher concentrations of IGF-I, IGFBP-3 and IGF-II, respectively in comparison with men with <secondary/high school education, and 15% lower concentrations of IGFBP-1 (Figure 5.12). Family history of prostate cancer was not associated with IGF or IGFBP concentrations (Figure 5.13).

5.3.9 Tests for heterogeneity and sensitivity analysis

Some heterogeneity was observed between studies in the associations of IGF and IGFBP concentrations with anthropometric and sociodemographic factors. When the associations were assessed separately for each study the direction was the same; the heterogeneity appeared to be due to differences in the magnitudes of associations (data not shown).

Similarly, heterogeneity was observed across studies in the associations between age and IGFBP-1, and BMI and IGFBP-2 by fasting status. Again, this heterogeneity appeared to be due to differences in the magnitudes of the associations (Figures 5.14 and 5.15).

The results remained broadly similar after restricting the dataset to: i) non-Hispanic white men; ii) ELISA only assays; and iii) excluding outlier values for IGF and IGFBP concentrations. Removing the adjustments for height made no material difference to the associations nor did after additional adjustment for smoking status and alcohol consumption.

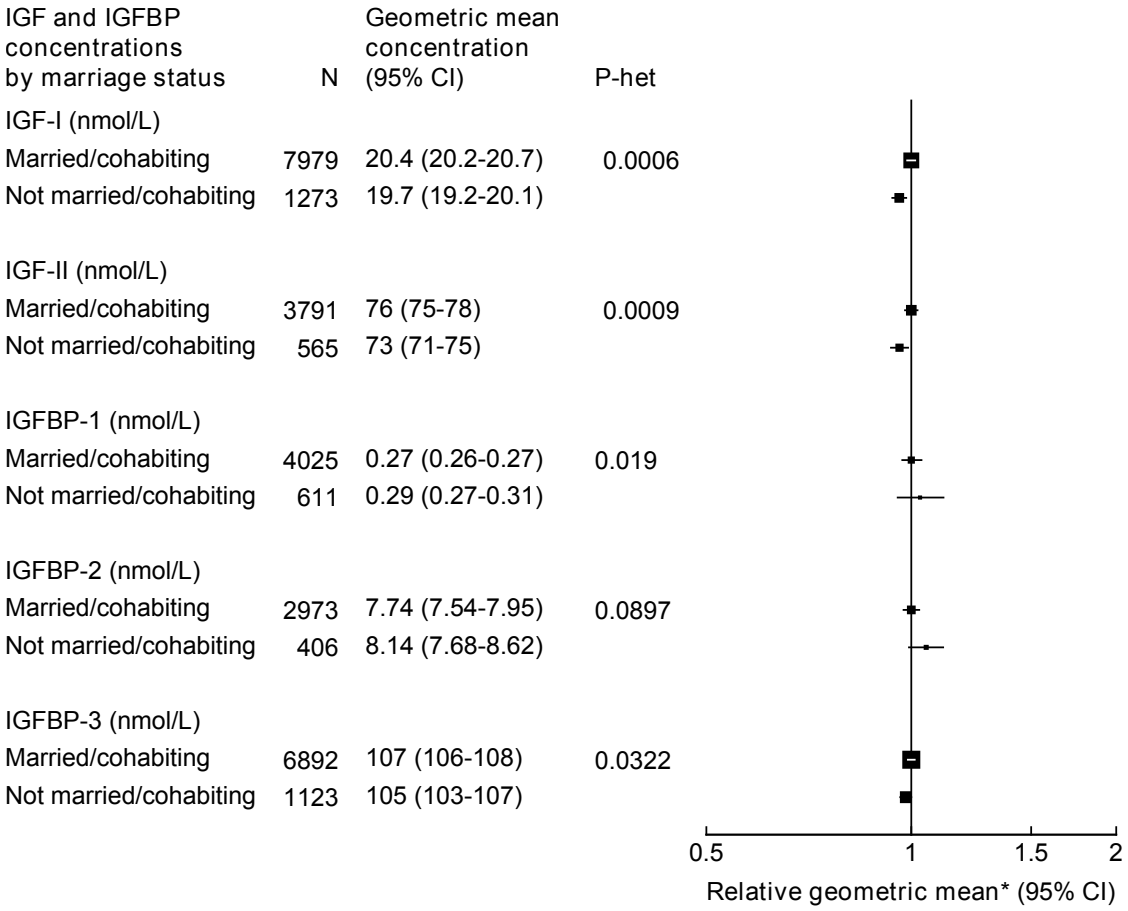


Figure 5.11: Relative geometric mean concentrations* of IGFs and IGFBPs in males by marriage status, adjusted for study, age, height and BMI.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. There was no statistically significant heterogeneity by study.
*relative to currently married/cohabiting.
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

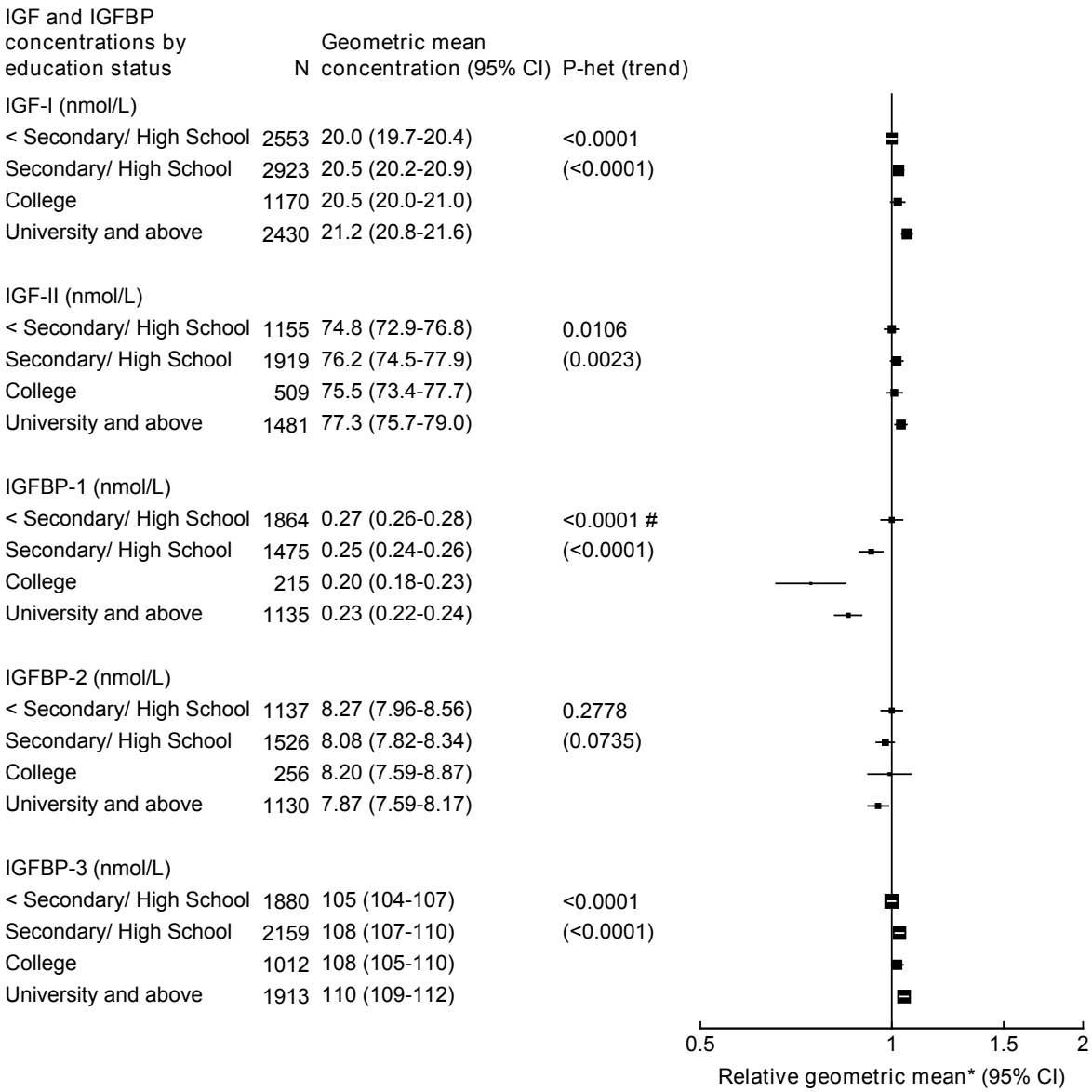


Figure 5.12: Relative geometric mean concentrations* of IGFs and IGFBPs in males by education status, adjusted for study, age, height and BMI.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. There was no statistically significant heterogeneity by study.
*relative to no secondary/high school education.
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

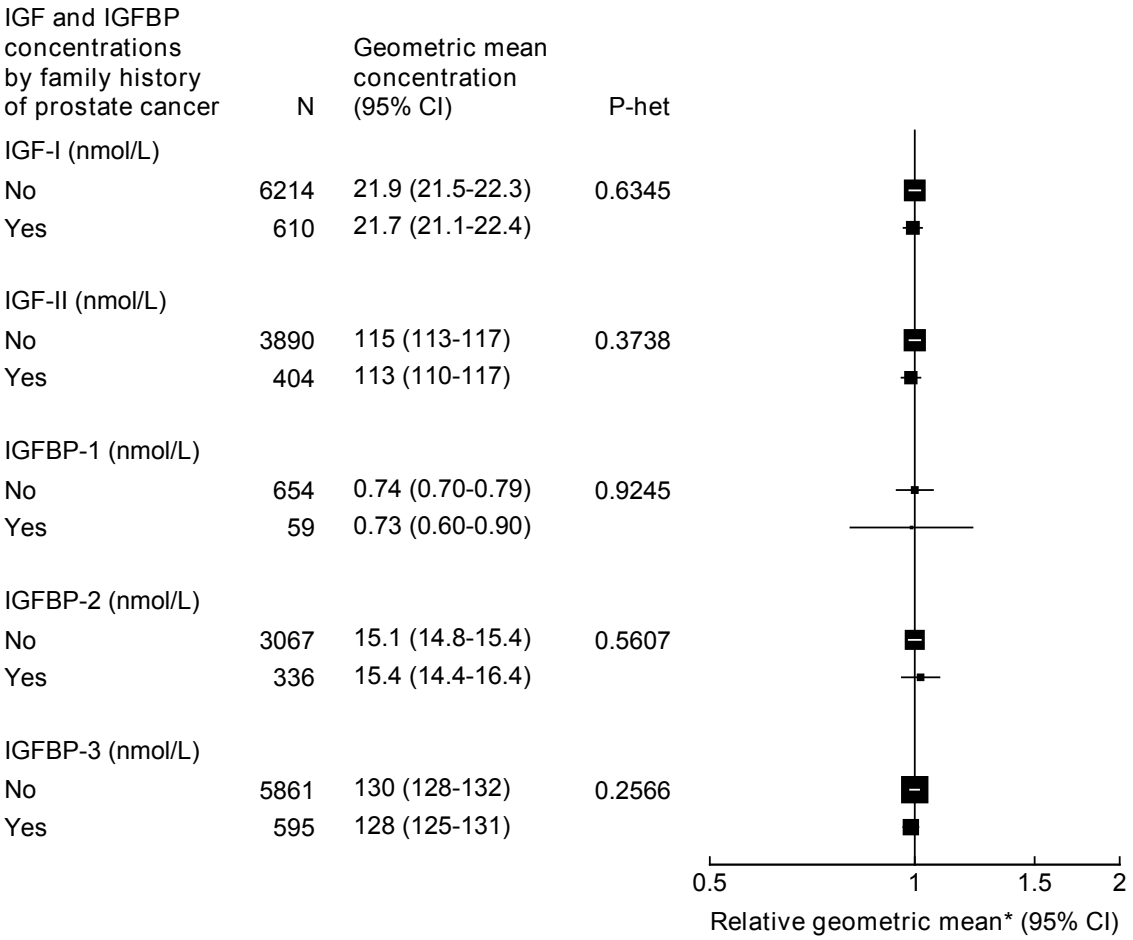


Figure 5.13: Relative geometric mean concentrations* of IGFs and IGFBPs in males by family history of prostate cancer, adjusted for study, age, height and BMI.
P for heterogeneity is the heterogeneity of means between categories, tested using the F test. There was no statistically significant heterogeneity by study.
*relative to no family history of prostate cancer.
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein.

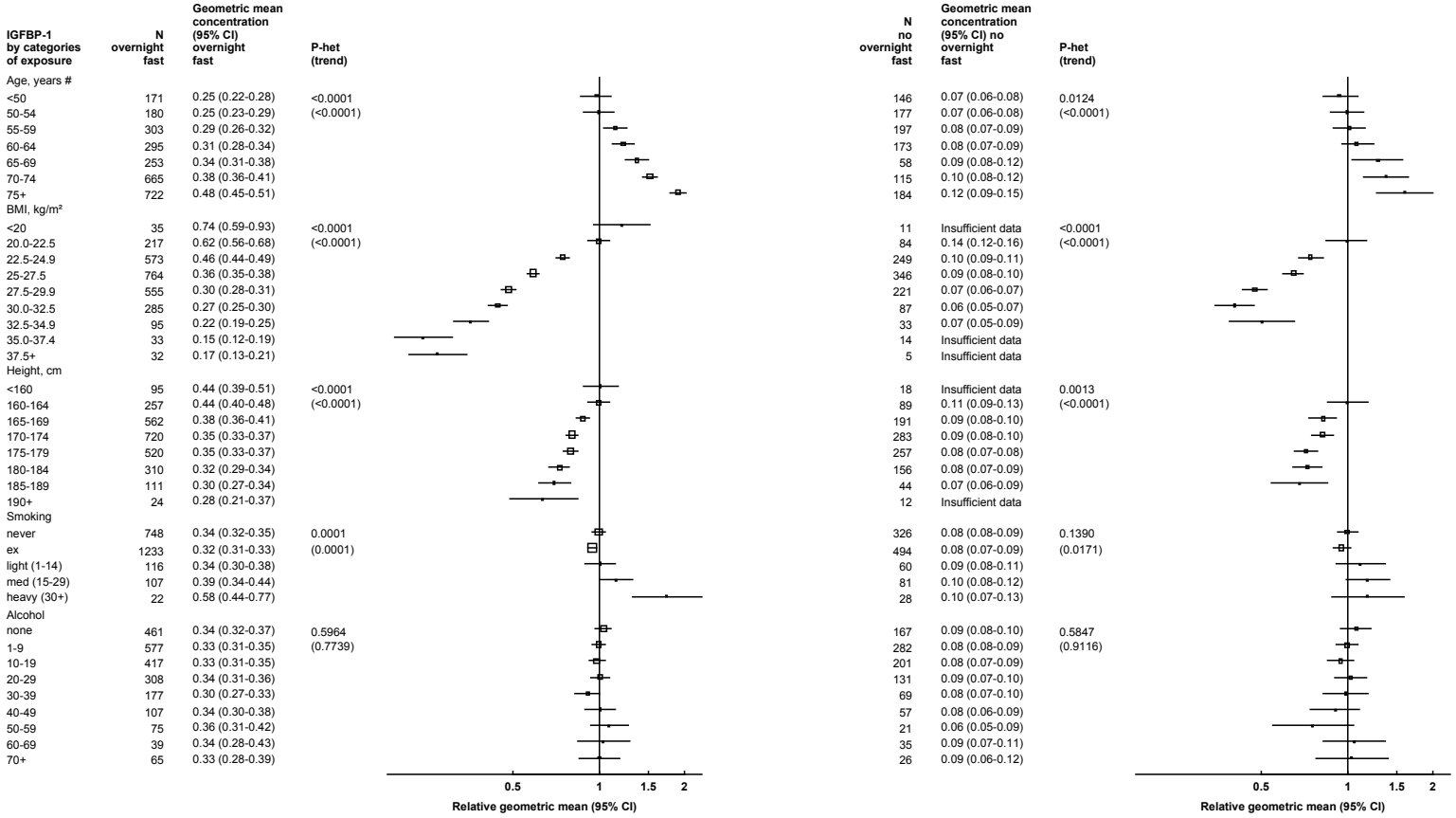


Figure 5.14: Relative geometric mean concentrations of IGFBP-1 by fasting status.
P for heterogeneity is the heterogeneity of means between categories, tested using the LR test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
#Significant heterogeneity by fasting status ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein; LR=likelihood ratio.

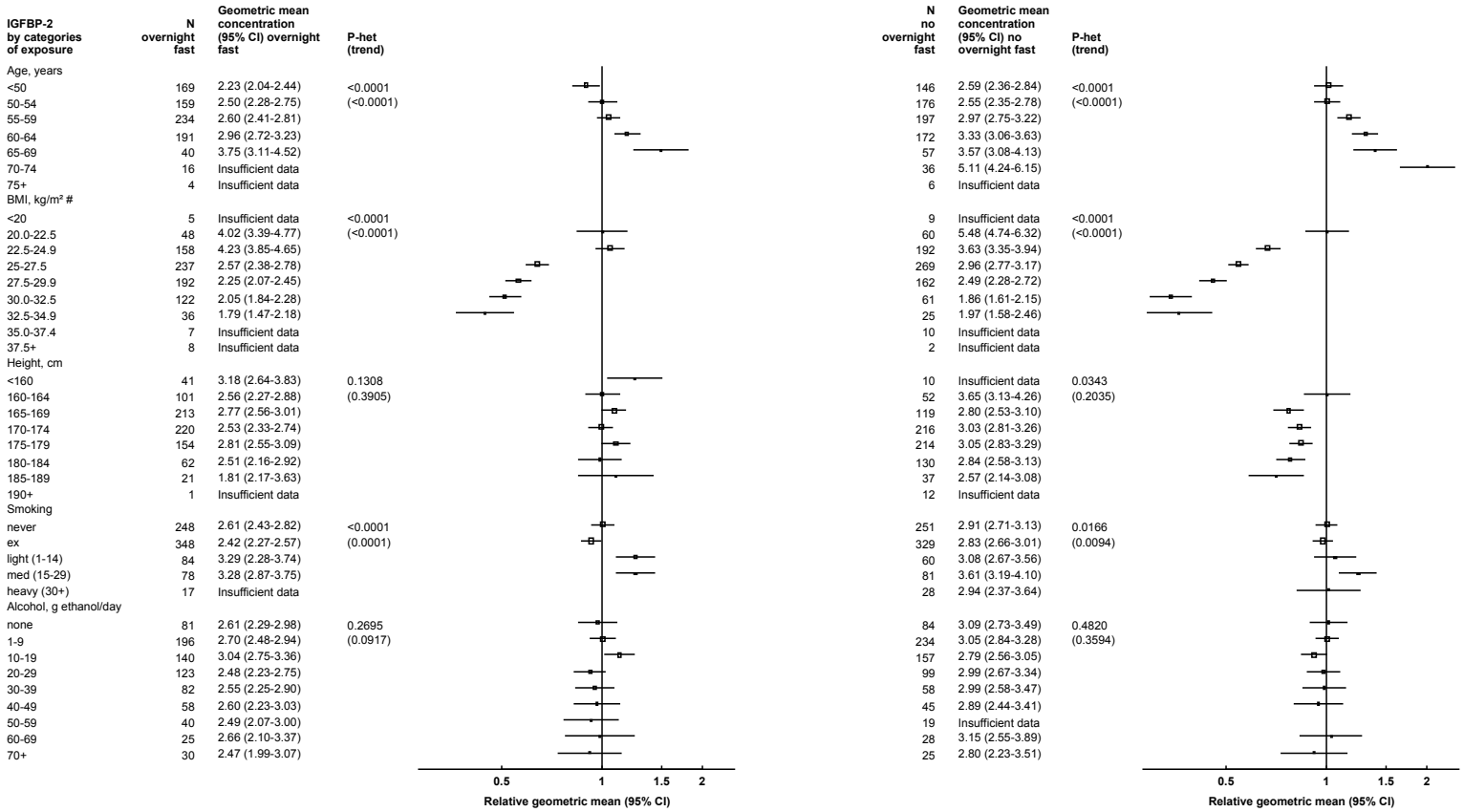


Figure 5.15: Relative geometric mean concentrations of IGFBP-2 by fasting status.
P for heterogeneity is the heterogeneity of means between categories, tested using the LR test. P for trend was calculated using the analysis of variance test, with categorical variables entered as linear values scored consecutively.
Significant heterogeneity by fasting status ($p<0.01$).
Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein; LR=likelihood ratio.

5.3.10 Further analyses of IGFs and sex hormones

Correlations between IGFs and sex hormones

Partial correlations between IGFs and sex hormones are displayed in Figure 5.16. After standardising for study and adjusting for age, BMI and height, IGF-I was negatively correlated with sex hormone binding globulin (SHBG) (r=-0.17) and positively correlated with free testosterone (r=0.11). IGF-II was negatively correlated with SHBG, testosterone and dihydrotestosterone (DHT) (r=-0.26, -0.14, and -0.13, respectively) and positively correlated with androstenedione (r=0.13). IGFBP-1 was positively correlated with SHBG, testosterone and DHT (r=0.22, 0.17 and 0.13, respectively). IGFBP-2 was positively correlated with SHBG, testosterone and estradiol (r=0.36, 0.24 and 0.11, respectively). IGFBP-3 was negatively correlated with SHBG (r=-0.27) and positively correlated with free testosterone (r=0.11).

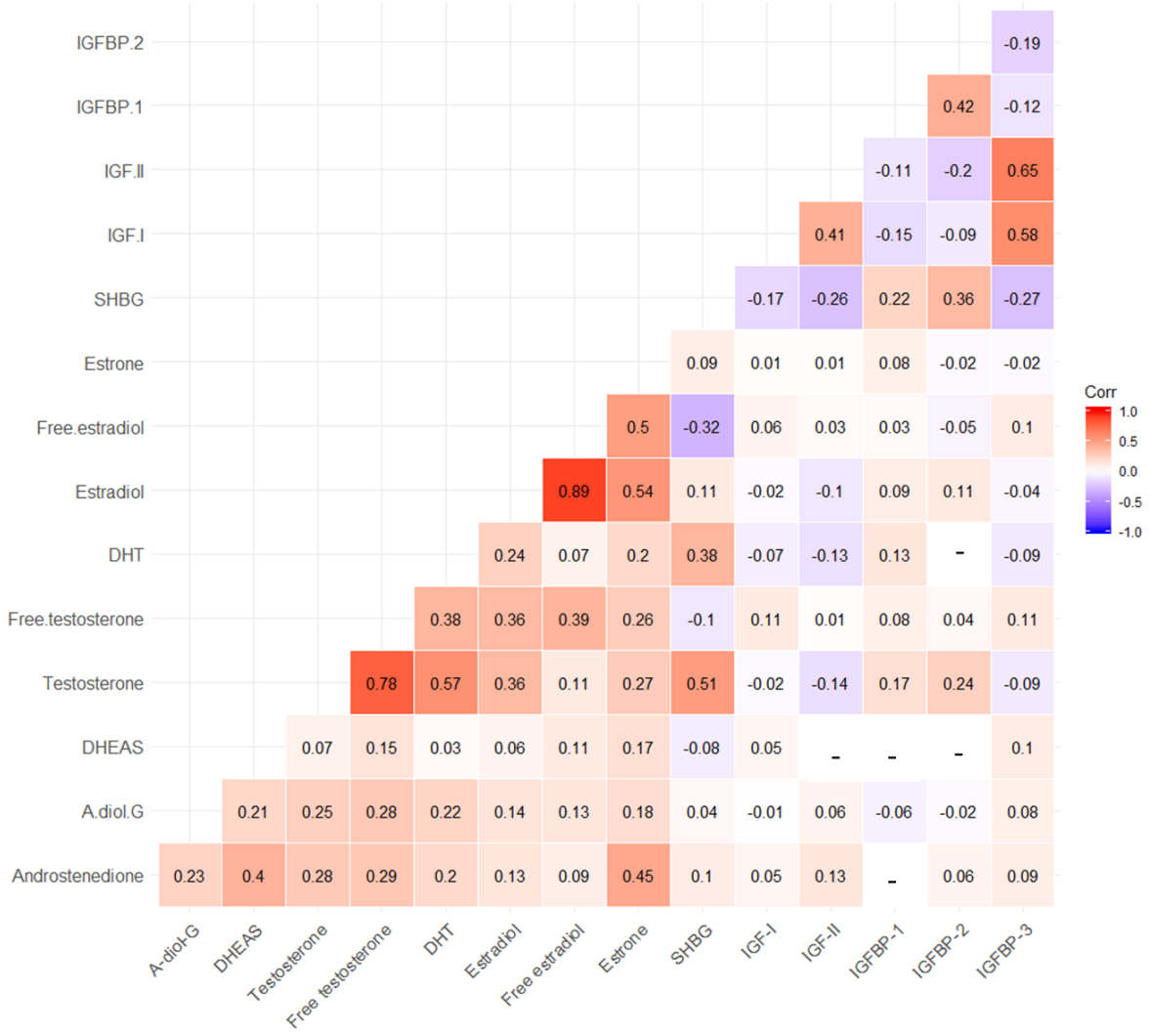


Figure 5.16: Partial correlation coefficients in control subjects between log-transformed sex hormones, SHBG, IGFs and IGFBPs.
Values are standardised by study and adjusted by age, BMI and height.
Number of observations between sex hormones and IGFs ranged from 7,469 (testosterone and IGF-I) to 188 (estrone and IGFBP-1).
Abbreviations: A-diol-G=androstanediol glucuronide; BMI=body mass index; DHEAS=dehydroepiandrosterone sulfate; DHT=dihydrotestosterone; IGF=insulin-like growth factor; IGFBP=insulin-like growth factor binding protein; SHBG=sex hormone binding globulin.

5.4 Discussion

This analysis of individual participant data from 23 studies, confirms strong associations of age, anthropometric factors and racial/ethnic groups with circulating IGFs and their binding proteins, and shows some novel associations, particularly with smoking and alcohol consumption. The large size and inclusion of different populations enabled us to assess associations across a wider range of exposure distributions than has previously been possible.

5.4.1 Correlations between hormones

In comparison with sex hormones, less is known about the possible determinants of the IGFs and their binding proteins. The liver is the predominant source of IGFs and IGFBPs. Hepatic IGF-I production is regulated by growth hormone (GH) [386]. IGFBP-3 concentrations are correlated with IGF-I, but hepatic production may be GH independent [387]. Similarly, IGF-II concentration is correlated with IGF-I, but the determinants of its production are unclear [388]. IGFBP-1 and IGFBP-2 concentrations may be partially regulated by insulin [389–391].

5.4.2 Age

These results showed that age is associated with circulating concentrations of all IGFs and their binding proteins. IGF-I, IGF-II and IGFBP-3 concentrations were lower in older men, which has been reported previously [392]. Secretion of GH declines with age, leading to a reduction in IGF-I [393]. IGFBP-1 and IGFBP-2 concentrations were higher in older men. Insulin may inhibit IGFBP-1 and IGFBP-2 production [389–391]. As men age, insulin sensitivity decreases [389, 394] and beta-cell (pancreatic cells which secret insulin) function is reduced [395], therefore ageing may reduce the insulin-related inhibition of IGFBP-1 and IGFBP-2.

5.4.3 Adiposity

BMI had modest inverse U-shaped association with circulating IGF-I, IGF-II and IGFBP-3 concentrations. Low energy consumption is associated with lower circulating IGF-I concentrations [176], therefore increasing energy consumption may explain the moderately higher IGF-I concentrations observed in men with BMI <20.0 kg/m^2 through to BMI 25.0-27.5 kg/m^2. Above this BMI, higher adiposity may impair liver function, reducing IGF-I secretion [396]. BMI also had strong inverse associations with IGFBP-1 and IGFBP-2, which may largely be driven by high insulin concentrations [389–391]. Changes in circulating IGFBP-1 concentrations may affect IGF-I bioavailability [397, 398], increasing the risk of neoplastic transformation [399, 400]; consequently, lower IGFBP-1 in men with higher adiposity may partly explain why men with high BMI may have an increased risk of aggressive prostate cancer [92]. Following additional adjustment for BMI, both waist circumference and WHR remained inversely associated with IGFBP-1 and IGFBP-2. This suggests that central adiposity may also be an important independent predictor of these concentrations.

5.4.4 Height

IGF-I plays a direct role in the regulation of growth in height [383]. In these analyses, height had modest positive associations with IGF-I and IGFBP-3, and a marked inverse association with IGFBP-1. Prospective studies consistently indicate that taller adult height is associated with an increased risk of prostate cancer [92], and it is possible that altered IGF-I bioavailability, due to lower IGFBP-1 concentrations [397, 398], may partially explain these associations.

5.4.5 Smoking

Smoking status was associated with differing circulating concentrations of IGFBPs. The inverse association between smoking and IGFBP-3 has been reported previously [401, 402]. Previous studies have not found an association between smoking and IGFBP-1 and IGFBP-2 [403, 404], which was observed in this analysis, but they did not stratify the associations by sex. The current analysis, however, did not show evidence of a linear trend when the association was tested excluding non-smokers. Smoking may be associated with an increased risk of prostate cancer progression, but there is currently no clear evidence that that smoking increases incident prostate cancer risk [405, 406],

5.4.6 Alcohol

Alcohol consumption was inversely associated with IGF-I and IGFBP-2 concentrations, and positively associated with IGF-II. Alcohol affects the liver by impaired insulin receptor binding and increased oxidative stress due to the toxicity of ethanol [407], which may in turn cause changes in IGF synthesis by the liver. To-date there are no studies which have reported an association between alcohol and IGF-II or IGFBP-2, although in mice hepatic IGF-II has been shown to increase after liver injury as part of the repair process [408]. Alcohol consumption is associated with short-term decreases in IGF-I and increases in IGFBP-1 concentrations, but concentrations normalise following alcohol metabolism [409, 410]. No data were available in this dataset regarding the men's time or quantity of last drink prior to blood sampling. There is some evidence that alcohol consumption may reduce prostate cancer risk and progression [411]; these results indicate that the lower IGF-I concentrations in men that consume alcohol may partially explain these associations.

5.4.7 Racial/ethnic group

Five studies had adequate representation of men across different racial/ethnic groups, all from the USA. Hispanics and African American men had lower IGF-II, IGFBP-1, IGFBP-2, and IGFBP-3 concentrations compared with non-Hispanic

whites. Hispanics also had lower IGF-I concentrations than white men [367, 412], whereas concentrations of IGF-I in African Americans were similar to those in non-Hispanic whites. Overall, Hispanics have a lower incidence of prostate cancer, in comparison with non-Hispanic whites, whilst African Americans have an increased incidence of prostate cancer [364, 365]. In Hispanics, the reduction in both IGF-I and IGFBPs may reduce the concentration of bioavailable IGF-I, whereas the lower concentrations of IGFBPs in African Americans might lead to an increase in bioavailable IGF-I. These differences may contribute to differences in prostate cancer risk observed between racial/ethnic groups.

5.4.8 Sociodemographic factors and health

Only modest differences in IGF-I and IGF-II concentrations were observed between married/cohabiting and single men, and there was no detectable difference by family history of prostate cancer. There was evidence that men with higher levels of education had different IGF and IGFBP concentrations. This has been observed elsewhere [413] and may be related to nutrition and/or health status [354, 414].

5.4.9 Heterogeneity by study

For several associations there was heterogeneity between the included studies. This was generally due to differences in the magnitudes of associations, rather than the direction. Heterogeneity may have been due to differences in sample populations, study design and measurement. Most studies used ELISA assays, although exact methods varied. Restricting the analysis to ELISA only assays did not materially affect the associations, but it is possible that assay method may contribute to some of the observed differences in the magnitudes of the associations [415, 416].

5.4.10 Correlations between IGFs and sex hormones

Most partial correlations between IGFs and sex hormones were moderate-to-weak. IGFs and IGFBPs consistently had the strongest correlations with SHBG. Like the IGFs, SHBG is primarily produced in the liver [161]. IGFs, as well as insulin, may

downregulate hepatic SHBG production [162], and higher SHBG concentrations may lead to the upregulation of estradiol and testosterone production by reducing free hormone concentrations, thus stimulating further hormone synthesis via negative feedback [164] (Chapter 2). Therefore, the moderate correlations observed between IGFs and sex hormones may be partially explained by SHBG.

5.4.11 Strengths and limitations

This analysis has a number of strengths. The large size of the dataset enabled the assessment of associations across a wider range of exposure distributions than previously possible, particularly for IGF-II, IGFBP-1 and IGFBP-2, which have been the subject of less research. This analysis also accounted for differences between studies to enable comparisons of the magnitudes of the associations relative to the different IGFs and their binding proteins.

The analysis also has some limitations. The data are cross-sectional and therefore it is not possible to infer whether the observed associations are causal or investigate the possibility of reverse causation. Assay quality varied between studies, which may lead to considerable differences in measured IGF-I concentrations [416]. This analysis was also based on single measurements, on the assumption that these measures are representative of participants' hormone concentrations over the medium-to-long term. Studies suggest that these hormones and binding proteins have moderate temporal reproducibility over a one-to-five year period [368, 380, 417, 418] (e.g. intraclass correlation coefficients for IGFs and IGFBPs r=0.63-0.87 over a mean 1.3 year period) [417], but the extent to which these measures are reproducible over the longer term is less clear. In this dataset, no data were available for some other potential determinants or confounders such as physical activity [419], diet [366, 420], co-morbidities (such as diabetes and kidney or liver disease) or medications [110, 421–423]. Fasting status was missing for 58% of IGFBP-1 and IGFBP-2 data.

5.5 Conclusion

In conclusion, this large collaborative analysis indicates that age, BMI, height and ethnic/racial group are strongly associated with circulating IGF and IGFBP concentrations, whilst smoking status and alcohol consumption have more modest associations. IGFs and IGFBPs might therefore account for some of the observed association of BMI, height and race/ethnicity with risk of prostate cancer. These findings will improve the understanding of how these factors may influence prostate cancer risk.

6

Analysis of circulating free testosterone and prostate cancer risk

Contents

6.1 Introduction 148
6.1.1 Context 148
6.1.2 Biological threshold 149
6.1.3 Aims 150
6.2 Subjects and methods 151
6.2.1 Data collection 151
6.2.2 Inclusion criteria 151
6.2.3 Study design and matching 151
6.2.4 Data processing 157
6.2.5 Statistical analysis 157
6.2.6 Tests for heterogeneity 158
6.2.7 Sensitivity analysis 159
6.2.8 Statistical software 159
6.3 Results 160
6.3.1 Associations between circulating sex hormones and prostate cancer risk 165
6.3.2 Heterogeneity by study 167
6.3.3 Heterogeneity by tumour and participant characteristics 167
6.3.4 Other assay measures 170
6.3.5 Sensitivity analysis 170
6.4 Discussion 172
6.4.1 Main findings 172
6.4.2 Biological plausibility 172
6.4.3 Heterogeneity by tumour grade 172
6.4.4 Detection bias 174

6.4.5 Co-morbidities 174
6.4.6 IGF-I at blood collection 175
6.4.7 Free hormone hypothesis 175
6.4.8 Strengths and limitations 176
6.5 Conclusion 178

6.1 Introduction

A previous analysis using the Endogenous Hormones, Nutritional Biomarkers Prostate Cancer Collaborative Group (EHNBPCCG) dataset did not find evidence of an association between any circulating sex hormone concentrations and prostate cancer risk [13]. Following this analysis further experimental and clinical evidence has emerged, including the results from two large randomised controlled trials, which were suggestive of a possible relationship between intra-prostatic androgen concentrations and prostate cancer risk. This association may be non-linear, as hypothesised by persistent intra-prostatic dihydrotestosterone (DHT) and/or saturation of the androgen receptors. Previous analyses may have been underpowered to detect this non-linear association. Since the analysis, published in 2008, the number of prostate cancer cases in this consortium has approximately doubled. Given the evidence summarised below and in Chapter 2, and that circulating free testosterone concentrations may be the primary determinant of intra-prostatic androgen concentrations, this chapter uses this updated EHNBPCCG dataset to re-examine the association between circulating free testosterone concentrations and subsequent prostate cancer risk in the context of a possible threshold association.

6.1.1 Context

In 1941 Huggins and Hodges demonstrated that reducing testosterone concentrations by castration led to prostate cancer regression, whilst exogenous testosterone administration led to prostate tumour progression [12, 211]. This discovery earned Charles Huggins a Nobel Prize in Physiology or Medicine in 1966.

Further clinical and experimental research also supports the role of testosterone in prostate cancer development. Nearly all metastatic prostate tumours overexpress the androgen receptor and androgen deprivation therapy is the mainstay treatment approach for many advanced stage prostate tumours [424]. Genome-wide association studies have also identified several single nucleotide polymorphisms that are associated with prostate cancer risk which code for signalling molecules that interact with, or are localised to, the androgen receptor [6, 7, 425]. Exogenous testosterone administration can induce prostate cancer in rats [224, 426].

Recently, two large randomised controlled trials investigated the effect of 5α-reductase inhibitors on prostate cancer risk [8–10]. These inhibitors effectively mimicked the effect of low circulating testosterone concentrations on the prostate by inhibiting the intra-prostatic conversion of testosterone to the more biologically active DHT, reducing intra-prostatic DHT concentrations by 80-90% [427]. Both trials reported a 23-25% relative reduction in prostate cancer risk in men who received the intervention [8–10].

Despite the apparently strong clinical and experimental evidence of an association between testosterone and prostate cancer pathogenesis, there is currently no consistent prospective epidemiological evidence of an association [13].

6.1.2 Biological threshold

The discrepancy between biological and clinical evidence and prospective epidemiological evidence of the role of testosterone in prostate cancer may be because the association is non-linear.

As described in Chapter 2, variations across the normal range of circulating testosterone concentrations may not lead to variations in the activation of the androgen receptor due to relatively constant intra-prostatic DHT concentrations and/or saturation of the androgen receptors [19, 196]. However, when the supply of testosterone to the prostate is abnormally low, androgen receptor signalling may decrease [194, 195]. Therefore, men with very low circulating testosterone concentrations may have a reduced risk of prostate cancer but, above these low

concentrations, further increases in circulating testosterone concentrations may not be associated with prostate cancer risk. Less than 2% of testosterone circulates unbound to carrier proteins or "free", and is able to pass out of the blood into the prostate tissue (Chapter 2) [159]; therefore, the focus of this analysis was on free testosterone. Prior epidemiological studies may have been under-powered to detect this non-linear association between low free testosterone and prostate cancer.

In 2008, the EHNBPCCG examined the associations between eight pre-diagnostic sex hormone concentrations (including free and total testosterone) and sex hormone binding globulin (SHBG), and prostate cancer risk. This was the largest epidemiological study to-date, using pooled individual-patient data from 18 studies including up to 3,886 cases and 6,438 controls, representing >95% of the worldwide data at the time of dataset closure [13]. This analysis examined the associations of circulating concentrations of these hormones and SHBG with prostate cancer risk in study-specific fifths and tenths and found no evidence of an association with any sex hormone, although there was evidence of an inverse association with circulating SHBG concentrations (risk ratio in the highest vs lowest fifth=0.86, 95% confidence interval (CI)=0.75-0.98, P_{trend}=0.01) [13].

Since this analysis, the EHNBPCCG dataset has been expanded to include almost double the number of prostate cancer cases and now comprises 20 prospective studies with a total of 6,933 cases and 12,088 matched controls with calculated free testosterone data. This large dataset may now provide enough power to detect a possible threshold association with prostate cancer.

6.1.3 Aims

Using this expanded EHNBPCCG dataset this analysis aims to test the hypothesis that men with low circulating free testosterone concentrations have a reduced risk of prostate cancer.

6.2 Subjects and methods

6.2.1 Data collection

The EHNBPCCG search strategy and pooling is discussed in greater detail in Chapter 3. In brief, collaborators were identified and requested to provide published or unpublished data on concentrations of endogenous hormones and/or nutritional biomarkers from blood samples collected prior to diagnosis of prostate cancer. Studies also provided individual participant data for baseline participant characteristics, subsequent prostate cancer status and tumour characteristics.

6.2.2 Inclusion criteria

Studies within the EHNBPCCG dataset were considered eligible for this analysis if they had prospective data on pre-diagnostic circulating concentrations of testosterone and SHBG, from which an estimate of free testosterone concentration could be calculated (see Chapter 2).

Although testosterone and SHBG data were available from the Prostate Testing for Cancer and Treatment (ProtecT) feasibility study, this study was not included in this analysis. This is because all men included in the trial underwent a prostate-specific antigen (PSA) test at study baseline. Those with PSA $\geqslant$3 ng/mL were offered a prostate biopsy [259]. Consequently, 50% of cases were diagnosed within 3 months of blood collection and 58% within the first year. Therefore, this study is considered cross-sectional rather than prospective.

Participants were excluded if they were known to be receiving androgen therapy at blood collection (n=2).

6.2.3 Study design and matching

20 prospective studies were identified with both total testosterone and SHBG measures, by dataset closure on August 31st, 2017. The studies were: Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC) [229], Baltimore Longitudinal Study of Aging (BLSA) [230, 231], Carotene and Retinol Efficacy Trial (CARET) [233], Child Health and Development Studies (CHDS) [234], European

Prospective Investigation into Cancer and Nutrition (EPIC) [237], EPIC Norfolk [238], Finnish Mobile Clinic (FMC) [240], Helsinki Heart Study (HHS) [243, 244], Health Professionals Follow-up Study (HPFS) [245], Health In Men Study (HIMS) [241, 242], Japan Collaborative Cohort Study (JACC) [246], Japan Public Health Center-based prospective study (JPHC) [249], Janus serum bank [244, 250], Melbourne Collaborative Cohort Study (MCCS) [252], Multiethnic Cohort (MEC) [254], Massachusetts Male Aging Study (MMAS) [255], Northern Sweden Health and Disease Cohort (NSHDC) [244], Prostate Cancer Prevention Trial (PCPT) [10], Physicians' Health Study (PHS) [256], and the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) [257].

Details of participant recruitment, study design, and case ascertainment are summarised in Table 6.1 and in the original publications. Most studies were matched case-control studies nested within either prospective cohort studies or randomised trials. Of the randomised trials, data were available for participants in both the intervention and placebo arm in ATBC, CARET, HHS, and PHS. Assay data were only collected in the screening arm of the PLCO trial, and the placebo arm in PCPT (as the intervention arm was designed to alter prostate hormone concentrations) [10]. Four studies were cohort or case-cohort analyses (BLSA, HIMS, MCCS and MMAS). To apply a consistent statistical approach across all studies, the cases from the case-cohort studies were matched to up to four participants who were free of prostate cancer at the age at diagnosis of the case on the basis of a minimal matching criteria (Table 6.2).

Each study individually obtained ethical approval; therefore, separate ethical approval for this secondary re-analysis of data were not necessary. Assay details and their performance characteristics are in Table 6.3.

Table 6.1: Sample population, recruitment and case ascertainment method by study.

Study	Sample population	Location	Recruitment period	Age at blood collection (years)	Prostate cancer ascertainment method
ATBC	Randomised trial of α-tocopherol and β-carotene among smokers	Finland	1985-1993	50-69	Cancer registry linkage, central review of medical records and specimens
BLSA	Prospective cohort study of the physiology of aging	USA	1958-onward	30-84	Self-report with medical record review
CARET	Randomised controlled trial of β-carotene and retinyl palmitate in heavy smokers and asbestos workers	USA	1987-1998	47-77	Self-report with medical record review
CHDS	Prospective cohort study of the husbands of pregnant women who were health plan members	USA	1959-1966	20-55	Cancer registry linkage
EPIC Phase 1	Population-based prospective cohort study	Europe	1991-2001	43-76	Cancer registry linkage; health insurance record linkage; Self-report with medical record review
EPIC Norfolk	Population-based prospective cohort study	UK	1991-1997	48-77	Cancer registry linkage
FMC	Population-based prospective cohort study	Finland	1966-1972	17-80	Cancer registry linkage
HHS NBSBWG	Randomised controlled trial of gemfibrozil	Finland	1981-1982	40-56	Cancer registry linkage
HIMS	Population-based randomized trial of screening for abdominal aortic aneurysms	Australia	1996-1999 (Blood collection 2001-2004)	71-87	Cancer registry linkage
HPFS	Cohort study of male health professionals	USA	1986	46-87	Self-report with medical record review
JACC	Population-based prospective cohort study of participants attending screening programs	Japan	1988-1990	57-85	Cancer registry linkage
JPHC	Population-based prospective cohort study	Japan	Cohort 1) 1989-1991 Cohort 2) 1993-1995	41-71	Active patient notification from major local hospitals in the study area and linkage with population-based cancer registries
Janus NBSBWG	Population-based prospective cohort study of participants in the Norwegian Regional Health Studies and Red Cross blood donors	Norway	1973-onward	33-61	Cancer registry linkage
MCCS	Prospective cohort study of inhabitants of the Melbourne metropolitan area	Australia	1990-1994	40-72	Cancer registry linkage
MEC	Ethnically diverse prospective cohort study	USA	1993-1996 (Blood collection 2001-2006)	48-85	Cancer registry linkage
MMAS	Prospective cohort study of the physiology of ageing	USA	1987-1989	41-70	Cancer registry linkage

Continued on next page

Table 6.1 – continued from previous page

Study	Sample population	Location	Recruitment period	Age at blood collection (years)	Prostate cancer ascertainment method
NSHDC	Combination of a population-based intervention study to decrease cardiovascular disease and a population-based monitoring study of cardiovascular disease	Sweden	1985-onward	39-61	Cancer registry linkage
PCPT	Randomised, placebo-controlled trial of finasteride and prostate cancer	USA	1994-1997	55-83	Diagnosed as part of trial protocol. Annual digital rectal examinations and PSA measurements. Biopsy if abnormal DRE or reported PSA level > 4.0 ng per. End-of-study prostate biopsy
PHS	Randomised trial of aspirin and β-carotene among physicians	USA	1982-onward	41-78	Self-report with medical record review
PLCO	Population-based randomised controlled multicentre trial of methods for early detection of cancer of the prostate, lung, colorectum and ovary	USA	1993-2001	55-74	Medical and pathology record review after screening and self-report with medical record review

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Heart Study; HIMS=Health In Men Study; HPFS=Health Professionals Follow-up Study; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NBSBWG=Nordic Biological Specimen Biobank Working Group; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; PSA=prostate-specific antigen.

Table 6.2: Criteria used by individual studies to match case patients and control subjects.

Study	Case: control ratio	Age at recruitment	Date of recruitment	Time of blood draw	Other matching criteria
ATBC	1:2	±1 y	±28 days		Trial intervention group, study centre
BLSA*	1:1	±6 mo			Follow-up time
CARET	1:1	±5 y	Month from enrolment to blood draw, year of randomisation	±2 h	Study centre, ethnicity
CHDS	1:2	±1 y			Ethnicity
EPIC Phase 1	1:1 except for Umea centre (1:2)	±6 mo		±1 h	Recruitment centre, time between blood draw and last food or drink consumption, follow-up time
EPIC-Norfolk	1:2	±3 y	±3 mo		Follow-up time
FMC	1:2	Nearest available	±1 mo		Municipality
HHS NBSBWG	1:4	±2 y	±2 mo		Study centre
HIMS*	1:4	±1 y			Fasting status, diabetes, controls must be 'alive and at risk' beyond the case's date of diagnosis. ±1 y date of blood collection
HPFS	1:1	Year of birth ±1 y	Exact year	Midnight–9 am; 9 am–12 pm; 12 pm–4 pm; and 4 pm–midnight	PSA test before blood draw, season, control subjects had at least one PSA test after the date of blood draw
JACC	1:3	As close as possible			Recruitment area
JPHC	1:2	±3 y	±60 days	±3 h	Municipality, duration of fasting at blood collection (±3 h)
Janus NBSBWG	1:4	±2 y	±6 mo		County of residence and Red Cross blood donor status (Oslo)
MCCS*	1:3	±5 y	±2 y		Assay batch, country of birth
MEC	1:2	±1 y	±6 mo	±2 h	Geographic site, ethnicity, fasting status (<6, 6-7, 8-9, 10+ hours)
MMAS*	1:4	±2 y	± 2 y		Ethnicity
NSHDC	1:2	±6 mo	±2 mo		County of residency
PCPT	1:1	As close as possible	As close as possible		PCPT treatment arm (placebo only). All non-white controls were sampled and then back filled with whites to achieve frequency matching on age and family history. Controls were required to have completed end of study biopsy procedure.
PHS	1:2	±1 y			Had not had a total or partial prostatectomy and smoking status
PLCO	1:1	±5 y	Exact year		Follow-up time

*Used a case–cohort design that was subsequently converted into nested case–control design.

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Heart Study; HIMS=Health In Men Study; HPFS=Health Professionals Follow-up Study; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NBSBWG=Nordic Biological Specimen Biobank Working Group; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

Table 6.3: Assay methods and geometric mean testosterone and SHBG concentrations by study.

		Testosterone (nmol/L)			SHBG (nmol/L)		
Study		Method	CV %	Geometric mean (95% CI)	Method	CV %	Geometric mean(95% CI)
ATBC	Case	E RIA	5.5†	20.1 (18.7-21.6)	IRMA	4.2†	82.6 (75.9-89.9)
	Control			20.5 (19.5-21.5)			85.3 (80.4-90.5)
BLSA	Case	NE RIA	3.3-6.4§	15.1 (14.0-16.2)	NE RIA	1.8-22§	76.2 (70.0-83.0)
	Control			15.3 (14.2-16.4)			81.1 (74.5-88.3)
CARET	Case	E RIA	1-12§	13.5 (12.9-14.1)	IRMA	2-7§	27.4 (26.0-28.9)
	Control			14.1 (13.5-14.7)			27.1 (25.8-28.6)
CHDS	Case	E RIA	9-11‡	22.0 (21.2-22.9)	IA	4-6‡	32.7 (31.2-34.1)
	Control			22.1 (21.5-22.7)			33 (31.9-34.1)
EPIC Phase 1	Case	NE RIA	10.8-14.8†	15.9 (15.4-16.4)	IRMA	7.7-12.2†	41.3 (39.8-42.9)
	Control			15.8 (15.3-16.3)			42.7 (41.2-44.3)
EPIC Norfolk	Case	ECIA	7.1-12.0§	15.7 (14.4-17.1)	ECIA	4.6-5.7§	41 (37.2-45.3)
	Control			15.2 (14.7-15.6)			42 (40.6-43.5)
FMC	Case	NE RIA	4.5-7.2‡	23.5 (22.1-24.9)	IMF	6.6-8.7‡	51.4 (47.9-55.2)
	Control			23.0 (22.0-24.0)			50.5 (47.9-53.2)
HHS NBSBWG	Case	NE RIA	5.5-13§	20.0 (18.4-21.7)	IMF	1.3-10.1§	50.5 (45.7-55.7)
	Control			19.7 (18.9-20.6)			50.2 (47.6-52.9)
HIMS	Case	LC-MS/MS	<6‡	12.3 (11.8-12.8)	IA	<7‡	38.9 (37.0-40.9)
	Control			12.1 (11.8-12.3)			39.6 (38.6-40.6)
HPFS	Case	ECIA	≤5.2†¶	15.5 (15.0-15.9)	IRMA	≤11.5†¶	56.9 (54.9-58.9)
	Control			15.4 (15.0-15.9)			56.9 (55.0-58.9)
JACC	Case	NE RIA	5-12*	16.1 (14.2-18.2)	IRMA	5.6-6.9*	40.4 (35.0-46.6)
	Control			15.7 (14.5-16.9)			43.7 (40.0-47.8)
Janus NBSBWG	Case	E RIA	5.5-13.0§	21.7 (20.9-22.4)	IMF	5-15§	46.6 (44.8-48.5)
	Control			22.4 (22.1-22.8)			48.2 (47.2-49.1)
JPHC	Case	ECIA	1-3§	15.4 (14.5-16.2)	IRMA	2-8§	46.0 (43.2-49.1)
	Control			15.7 (15.1-16.3)			47.5 (45.4-49.7)
MCCS	Case	ECIA	1.6*	15.3 (14.8-15.8)	IA	6*	37.3 (35.9-38.7)
	Control			15.4 (15.0-15.8)			36.6 (35.6-37.6)
MEC	Case	E RIA	3.5†	18.2 (17.6-18.9)	ECIA	3†	35.3 (33.9-36.8)
	Control			18.3 (17.8-18.7)			36 (34.9-37.0)
MMAS	Case	E RIA	4.6-7.2§	16.6 (15.6-17.6)	Filtration assay	8-10.9§	30.3 (28.2-32.5)
	Control			16.4 (15.9-16.9)			30.4 (29.4-31.5)
NSHDC	Case	N/S		20.4 (19.6-21.2)	IRMA	N/S	42.2 (40.3-44.1)
	Control			19.5 (18.8-20.3)			40.9 (39.0-42.8)
PCPT	Case	ECIA	7.6-11.9§	12.6 (12.3-12.9)	ECIA	5.2-12.2§	36.6 (35.6-37.7)
	Control			12.4 (12.2-12.7)			36.6 (35.6-37.6)
PHS	Case	NE RIA	8.7†	16.0 (15.2-16.9)	IRMA	8.9†	19.8 (18.6-21.0)
	Control			16.0 (15.4-16.6)			20.6 (19.6-21.6)
PLCO	Case	NE RIA	14*	16.2 (15.7-16.6)	IRMA	18*	42.9 (41.5-44.4)
	Control			15.9 (15.5-16.3)			44.0 (42.7-45.3)

*Not specified.
†Intra-assay.
‡Inter-assay.
§Intra-and inter-assay range.
¶Three separate assay phases.

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CV=coefficient of variation; E RIA=extraction radioimmunoassay; ECIA=electrochemiluminescence immunoassay; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Heart Study; HIMS=Health In Men Study; HPFS=Health Professionals Follow-Up Study; IA=immunoassay; IMF=immunofluorometry; IRMA=immunoradiometric assay; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; LC-MS/MS=liquid chromatography-tandem mass spectrometry; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NBSBWG=Nordic Biological Specimen Biobank Working Group; NE RIA=non-extraction radioimmunoassay; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

6.2.4 Data processing

Circulating free testosterone concentrations were estimated from total testosterone and SHBG concentrations using the law of mass action (see Chapter 4) [163, 302], assuming a constant albumin concentration of 43 g/L.

Prostate cancer cases were defined as early stage if they were tumour–node–metastasis (TNM) stage $\leqslant$T2 with no reported lymph node involvement or metastases, or stage I-II, or equivalent (i.e. a tumour that does not extend beyond the prostate capsule); advanced stage if they were TNM stage T3 or T4 and/or N1+ and/or M1, stage III–IV, or equivalent (i.e. a tumour extending beyond the prostate capsule and/or lymph node involvement and/or distant metastases); or "unknown" otherwise. Aggressive disease was categorised as "no" for TNM stage $\leqslant$T3 with no reported lymph node involvement or metastases or equivalent, "yes" for TNM stage T4 and/or N1+ and/or M1 and/or stage IV disease or death from prostate cancer, or "unknown" otherwise. Prostate cancer was defined as low-intermediate grade if the Gleason score was <8 or equivalent (i.e. extent of differentiation good, moderate, or poor), high-grade if the Gleason score was $\geqslant$8 or equivalent (i.e. undifferentiated), or grade "unknown" otherwise.

6.2.5 Statistical analysis

Free testosterone concentration was categorised into study-specific tenths, with the 1st tenth defined as the lowest concentration and the 10th as the highest. Cut-points defined by the distribution in control participants. Study-specific cut-points were used to allow for any systematic differences between the studies in assay methods and blood sample types [428]. These differences mean that absolute clinical cut-points cannot be used, but study-specific categorisation of concentrations will allow for relative concentrations to be determined within studies [428, 429]. In the HPFS, assays were carried out in three distinct phases; therefore phase-specific tenths were calculated within this study.

Estimates of prostate cancer risk were calculated using logistic regression conditioned on the matching variables and adjusted for: age at blood collection

(continuous), body mass index (BMI) (<25, 25.0-27.4, 27.5-29.9, ⩾30.0 kg/m^2, unknown (6%)), height (<170.0, 170.0-174.9, 175.0-179.9, ⩾180.0 cm, unknown (6%)), usual alcohol intake (none, 1-9, 10-19, 20-39, ⩾40 g ethanol/day, unknown (32%)), smoking status (never, former, current, unknown (10%)), currently married/cohabiting (yes, no, unknown (23%)), education status (<secondary/high school, secondary/high school, university, unknown (24%)). These categories were selected *a priori* based on previous evidence of associations with prostate cancer risk reported in the literature, with cut-points based on measurement availability and data distributions. The highest tenth was used as the reference category.

To explore the association with greater power, these tenths were also grouped (1^{st}, 2^{nd}-4^{th}, 5^{th}-7^{th}, 8^{th}-10^{th}), with the 8^{th}-10^{th} tenths combined as the reference category. Then, in all further analyses, the 2^{nd}-10^{th} tenths were combined and used as the reference category. Where more than two categories of exposure were compared, the variance of the log risk in each all groups was calculated (from the variances and covariances of the log risk in all groups except the reference group) and used to obtain group-specific 95% confidence intervals (CIs), which facilitate comparisons across different exposure categories [430].

6.2.6 Tests for heterogeneity

Heterogeneity in the associations of free testosterone with prostate cancer risk among studies was assessed by comparing the χ^2 values for models with and without a study*analyte interaction term. This was tested between all studies as well as between studies which included organised screening in their study design (PCPT and PLCO) and those that did not.

Heterogeneity in the associations of free testosterone with prostate cancer risk was tested by: grade at diagnosis (low, high), stage at diagnosis (localised, advanced), aggressiveness of disease (not aggressive, aggressive), age at diagnosis (<65, 65+ years), years from blood collection to diagnosis (<5, 5+ years), age at blood collection (<65, 65+ years), date of diagnosis (pre 1990, 1990 onwards (as approximates of pre and post PSA era)), PSA at blood collection (<2 ng/mL, 2+ ng/mL),

educational attainment status (no university degree, university degree), BMI (<30, 30+ kg/m^2), racial/ethnic group (non-Hispanic white, African American/Caribbean, East Asian), smoking status (former/never, current), alcohol consumption (<10, 10+ g ethanol/day), currently married/cohabiting (yes, no), IGF-I concentration at blood collection(<study-specific median, study-specific median+), time of blood collection (before 12:00, after 12:00), and type of testosterone assay (extracted, not extracted).

Tests for heterogeneity for case-defined factors, in which controls in each matched set were assigned to the category of their matched case, were obtained by fitting separate models for each subgroup and assuming independence of the odds ratios (ORs) using a χ^2 test, which is analogous to a meta-analysis. Tests for heterogeneity for non-case defined factors were assessed with χ^2 tests of interaction between subgroups and the binary variable.

6.2.7 Sensitivity analysis

PSA and IGF-I concentrations at blood collection were available for subsets of participants. To investigate this further, geometric mean concentrations of IGF-I and PSA were calculated in the control participants, adjusted for age and BMI by study-specific tenths of free testosterone concentration. The relationship between low free testosterone and prostate cancer risk were also examined in these subsets before and after further adjustment for these variables (log-transformed PSA (continuous) and study-specific fifths of IGF-I concentrations).

The analyses were also repeated using an alternative method to calculate free testosterone concentration (see Appendix A).

6.2.8 Statistical software

All statistical tests were two-sided, with the statistical significance threshold set at $p<0.05$. Data analysis was carried out using Stata Statistical Software release 14.1 (Stata Corp., College Station, TX, USA), and figures were created in R version 3.2.3.

6.3 Results

A total of 20 studies, comprising 6,933 cases and 12,088 controls, were eligible for this analysis. Mean age at blood collection in each study ranged from 33.8 to 76.2 years (overall mean=59.8 years, standard deviation (SD)=11.5 years) and the period of blood collection ranged from 1958 to 2004. Study participants were predominantly non-Hispanic white (82.0%). The average time from blood collection to diagnosis was 6.8 years, the average age at diagnosis was 67.9 years (SD=7.2), and most cases were diagnosed between 1995 and 1999 (38.9%). Prostate cancers were mostly localised (55.2%) and/or low-grade (68.3%) (Table 6.4). The free testosterone concentration decile cut-points used for each study are shown in Table 6.5. Circulating free testosterone cut-points ranged between 100-300 pmol/L, in the BLSA and CHDS, respectively. Men in the lowest study-specific tenth of free testosterone were older, had a higher mean BMI, and lower PSA at blood collection than men with higher free testosterone concentrations (Table 6.6).

Geometric mean total testosterone and SHBG concentrations in controls are displayed in Figure 6.1, and the distribution of men in the lowest study-specific tenth of free testosterone by study-specific tenths of total testosterone and SHBG concentrations are displayed in Table 6.7. Collectively, Figure 6.1 and Table 6.7 suggest that low total testosterone concentrations may be a stronger predictor of low free testosterone than higher SHBG concentrations.

Table 6.4: Characteristics of patients diagnosed with prostate cancer by study*.

Study	N	Years from blood collection to diagnosis, %		Age at diagnosis, %		Year of diagnosis, %		Stage of disease†, %			Aggressive disease‡, %			Grade of disease§, %		
		<5	5+	<65	65+	<1995	1995+	Localised*	Advanced*	N/k	No*	Yes*	N/k	Low*	High*	N/k
ATBC	116	72.4	27.6	48.3	51.7	100	0	61.2	38.8	0	72.4	27.6	0	80.4	19.6	3.4
BLSA	112	15.2	84.8	18.8	81.3	67.9	32.1	79.2	20.8	52.7	64.3	35.7	37.5	86.5	13.5	20.5
CARET	298	82.9	17.1	37.6	62.4	45.3	54.7	68.7	31.3	28.2	90.7	9.3	28.2	88.7	11.3	8.1
CHDS	322	0	100	45.7	54.3	44.1	55.9	80.1	19.9	11.2	85.4	14.6	10.9	98.7	1.3	5.9
EPIC Phase 1	490	78.2	21.8	53.5	46.5	1.2	98.8	70.2	29.8	32.9	66.8	33.2	25.7	88.5	11.5	25.3
EPIC-Norfolk	76	13.2	86.8	7.9	92.1	0	100	75.7	24.3	7.9	58.9	41.1	3.9	71.4	28.6	90.8
FMC	166	13.3	86.7	22.3	77.7	100	0	N/A	N/A	100	N/A	N/A	100	N/A	N/A	100
HHS																
NBSBWG	84	7.1	92.9	66.7	33.3	76.2	23.8	62.1	37.9	21.4	62.1	37.9	21.4	N/A	N/A	100
HIMS	319	66.1	33.9	0	100	0	100	N/A	N/A	100	0	100	88.7	N/A	N/A	100
HPFS	682	81.7	18.3	27	73	11.3	88.7	82.9	17.1	43.3	96.4	3.6	43.3	89.9	10.1	9.7
JACC	40	45	55	7.5	92.5	55	45	N/A	N/A	100	N/A	N/A	100	N/A	N/A	100
JPHC	201	19.4	80.6	18.4	81.6	3.5	96.5	71.5	28.5	24.9	70.1	29.9	23.4	75.8	24.2	69.2
Janus																
NBSBWG	491	2.6	97.4	58.9	41.1	81.3	18.7	N/A	N/A	100	N/A	N/A	100	N/A	N/A	100
MCCS	548	41.2	58.8	30.1	69.9	15.3	84.7	90.6	9.4	1.5	98.3	1.7	1.5	86	14	2
MEC	463	93.5	6.5	19.9	80.1	0	100	N/A	N/A	100	0	100	94.4	99.8	0.2	4.5
MMAS	163	19.6	80.4	30.7	69.3	28.8	71.2	N/A	N/A	100	0	100	98.2	100	0	47.2
NSHDC	384	36.7	63.3	55.2	44.8	4.7	95.3	80.3	19.7	1	87.6	12.4	1	90.3	9.7	17.2
PCPT	1,032	23.1	76.9	22.7	77.3	0.4	99.6	98.3	1.7	2.6	99.1	0.9	2.6	95.1	4.9	2.4
PHS	219	30.1	69.9	32.4	67.6	100	0	78.4	21.6	2.7	69.5	30.5	2.7	91.1	8.9	7.8
PLCO	727	89.1	10.9	25.3	74.7	0	100	88.8	11.2	0	94.8	5.2	0	95	5	0.1

*Percentage value is for cases with an available free testosterone measurement.

† Stage of disease was defined as follows: localised if TNM was T2 or lower with no reported lymph node involvement or metastases, stage II or lower, or equivalent (i.e. a tumour that does not extend beyond the prostate capsule); advanced if TNM stage was T3 or T4 and/or N1+ and/or M1, stage III or IV, equivalent (i.e. a tumour extending beyond the prostate capsule and/or lymph node involvement and/or distant metastases), or unknown.

‡Aggressive disease was defined as T4 and/or N1+ and/or M1, or stage IV disease and/or death from prostate cancer. Overall, 4,661 (67%) of case patients had data on disease aggressiveness.

§ Histological grade was categorised as low-intermediate grade (Gleason sum <8 or cases coded as well, moderately, or poorly differentiated), high-grade (Gleason sum 8+ or cases coded as undifferentiated), or unknown.

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Heart Study; HIMS=Health In Men Study; HPFS=Health Professionals Follow-up Study; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NBSBWG=Nordic Biological Specimen Biobank Working Group; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

Table 6.5: Geometric mean free testosterone concentrations by study-specific tenths (1^{st} vs 2^{nd}-10^{th}) and 1^{st} decile cut-point.

	Geometric mean (95% CI) (pmol/L)		1^{st} decile cut-point (pmol/L)
	1^{st} tenth	2^{nd}-10^{th} tenth	
ATBC	133 (118-152)	235 (227-243)	161
BLSA	83 (69-100)	171 (165-178)	100
CARET	171 (156-189)	347 (339-356)	218
CHDS	244 (218-273)	421 (409-432)	300
EPIC Phase 1	127 (117-138)	304 (298-310)	154
EPIC- Norfolk	159 (134-189)	322 (309-335)	197
FMC	176 (157-198)	406 (395-418)	240
HHS NBSBWG	170 (149-194)	317 (307-328)	210
HIMS	97 (90-104)	233 (230-237)	147
HPFS	96 (89-103)	225 (222-229)	179
JACC	163 (131-204)	302 (287-318)	181
JPHC	130 (116-146)	266 (259-272)	169
Janus NBSBWG	178 (166-191)	368 (362-374)	230
MCCS	150 (140-160)	318 (313-322)	202
MEC	155 (143-167)	430 (422-437)	236
MMAS	151 (139-165)	387 (379-395)	210
NSHDC	168 (152-186)	386 (377-394)	235
PCPT	132 (124-139)	259 (256-263)	160
PHS	222 (200-246)	444 (433-455)	265
PLCO	133 (124-142)	311 (306-316)	170

Abbreviations: ATBC=The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BLSA=The Baltimore Longitudinal Study of Aging; CARET=The Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; CI=confidence interval; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Heart Study; HIMS=Health In Men Study; HPFS=Health Professionals Follow-up Study; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NBSBWG=Nordic Biological Specimen Biobank Working Group; NSHDC=Northern Sweden Health and Disease Cohort; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

Table 6.6: Characteristics of prostate cancer cases and controls by study-specific tenths of calculated free testosterone concentration.

	Cases		Controls	
	1^{st} tenth	2^{nd}-10^{th} tenth	1^{st} tenth	2^{nd}-10^{th} tenth
Age, years				
Mean (SD)	62.8 (10.2)	60.5 (10.6)	60.8 (12.3)	59.2 (12.0)
Height, cm				
Mean (SD)	175.9 (7.6)	175.1 (7.7)	174.6 (8.2)	174.5 (7.7)
BMI, kg/m^2				
Mean (SD)	27.5 (4.6)	26.4 (3.6)	27.1 (4.8)	26.3 (3.7)
PSA, ng/mL				
Mean (SD)	6.63 (32.8)	7.59 (137.0)	1.30 (1.45)	1.54 (2.7)
Smoking Status (%)				
Never	29.5	32.8	27.2	27.3
Former	48.4	41.8	39.1	38.4
Current	16.0	18.9	23.9	23.2
Unknown	6.1	6.5	9.8	11.1
Alcohol consumption, g/d				
Non-drinkers	20.0	18.1	18.0	16.1
<10	27.0	25.7	20.1	20.7
10+	32.1	33.3	29.0	29.8
Unknown	20.9	22.9	32.8	33.4
Ethnicity (%)				
Non-Hispanic white	84.0	84.0	80.9	80.8
African American	6.8	5.8	7.3	6.2
East Asian	3.0	5.1	5.0	5.9
Other	1.8	1.9	1.8	2.4
Unknown	4.4	3.2	5.0	4.6
Currently married/ cohabiting (%)				
Yes	75.6	73.5	65.8	65.8
No	10.9	10.9	10.4	10.1
Unknown	13.5	15.6	23.8	24.1

Abbreviations: BMI=body mass index; PSA=prostate-specific antigen; SD=standard deviation.

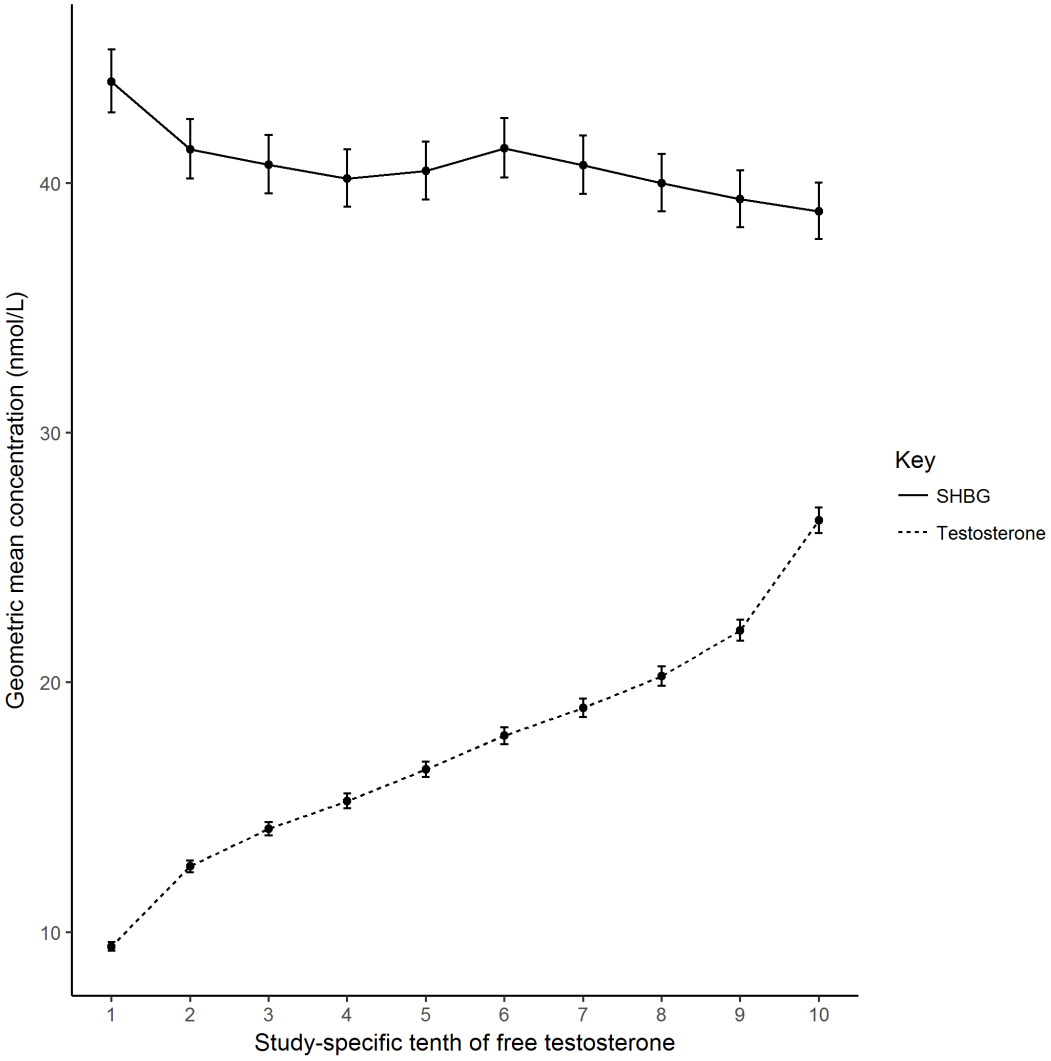


Figure 6.1: Geometric mean total testosterone and SHBG concentrations in controls by study-specific tenths of free testosterone.
Abbreviations: CI=confidence interval; SHBG=sex hormone binding globulin.

Table 6.7: Low free testosterone by study-specific tenths of total testosterone and SHBG.

Study-specific tenth		Percentage in 1st tenth of free testosterone		
		Overall	Cases	Controls
Testosterone	1	59.1	58.8	59.3
	2	18.6	18.1	18.8
	3	8.3	9.1	7.9
	4	4.4	4.9	4.2
	5	3.6	3.3	3.8
	6	2.4	2.8	2.2
	7	1.2	1.1	1.3
	8	1.2	1.1	1.2
	9	0.7	0.7	0.7
	10	0.5	0.2	0.6
SHBG	1	11.5	11.2	11.6
	2	10.0	10.4	9.8
	3	8.6	9.7	8.1
	4	8.1	8.1	8.1
	5	8.4	8.3	8.5
	6	7.6	8.6	7.1
	7	8.7	8.3	8.9
	8	9.3	8.8	9.6
	9	9.9	10.7	9.5
	10	17.9	16.1	18.7

Abbreviations: SHBG=sex hormone binding globulin.

6.3.1 Associations between circulating sex hormones and prostate cancer risk

Figure 6.2 shows the associations of free testosterone, total testosterone and SHBG concentrations with overall prostate cancer risk. Men in the lowest tenth of free testosterone had a lower risk of prostate cancer compared with men in any other tenth of the distribution in the adjusted model (Figure 6.2). The tenths were next combined into a smaller number of categories (1st, 2nd-4th, 5th-7th, 8th-10th); here, men in the lowest tenth had a 23% lower risk compared with men in the 8th-10th tenth group (Figure 6.3). When categories 2nd-10th were combined, the risk estimate remained very similar (OR for 1st versus 2nd-10th category=0.77, 95% CI 0.69-0.86; P<0.001).

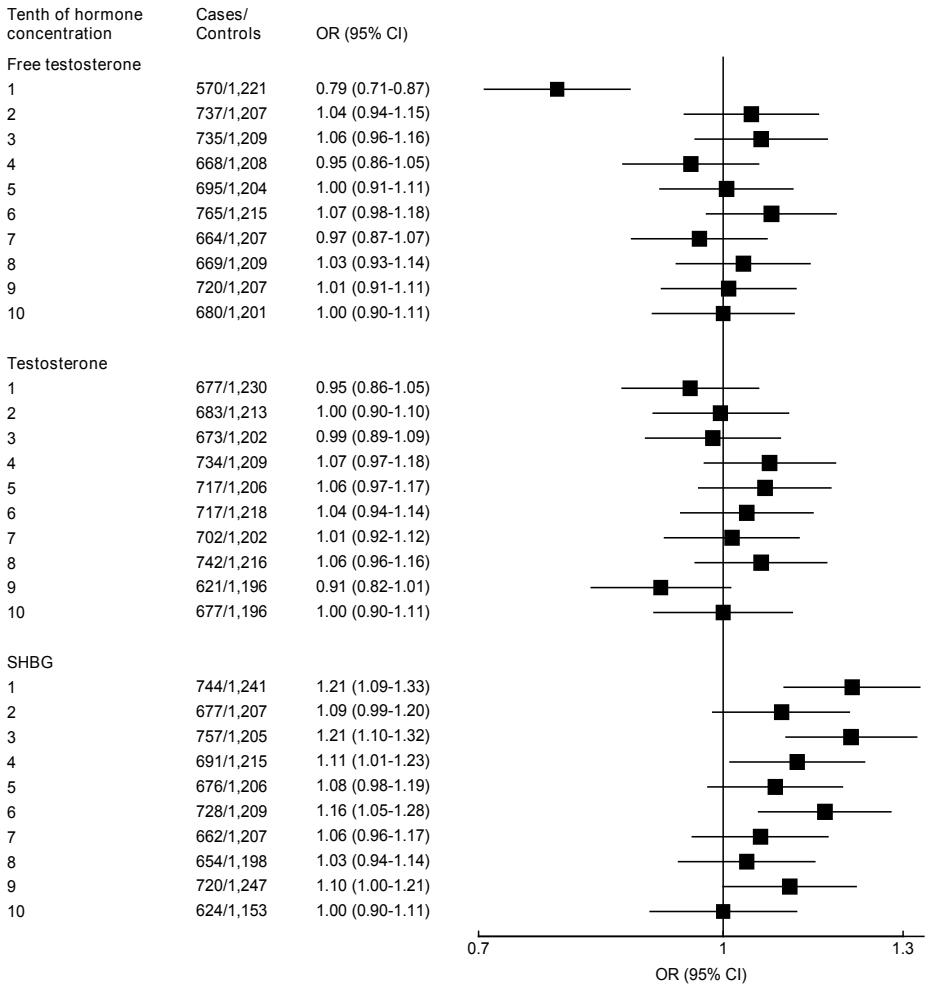


Figure 6.2: Odds ratio (95% CIs) of prostate cancer risk by study-specific tenths of hormone concentrations.

Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, and education status. The position of each square indicates the magnitude of the relative risk, and the area of the square is proportional to the amount of statistical information available (inverse of the variance of the logarithm of the relative risk). The length of the horizontal line through the square indicates the 95% confidence interval.

Abbreviations: BMI=body mass index; CI= confidence interval; OR=odds ratio; SHBG=sex hormone–binding globulin.

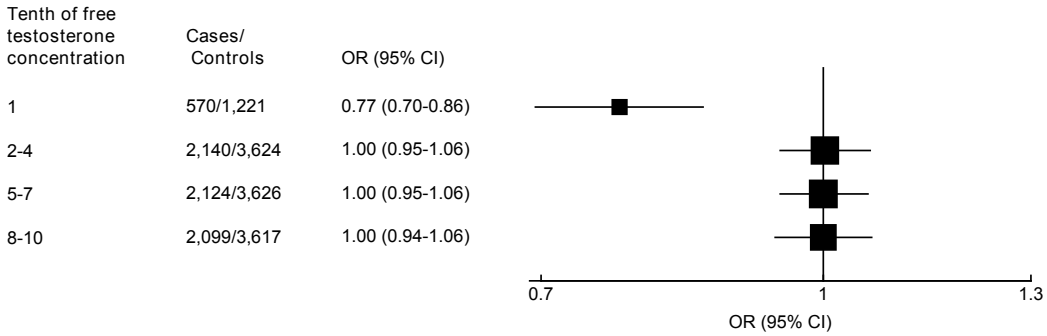


Figure 6.3: Odds ratio (95% CIs) of prostate cancer risk by study-specific tenths free testosterone concentrations.
Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, and education status.
Abbreviations: BMI=body mass index; CI= confidence interval; OR=odds ratio.

6.3.2 Heterogeneity by study

There was no evidence of heterogeneity in the association of low free testosterone and prostate cancer risk between studies (χ^2_{19}=18.0; P=0.53) (Figure 6.4). Two studies (PCPT and PLCO) included organised prostate cancer screening (25% of case participants), but there was no evidence of heterogeneity in the associations between the studies which did and did not include organised screening (χ^2_1=0.73; P=0.39).

6.3.3 Heterogeneity by tumour and participant characteristics

There was evidence of heterogeneity in the associations with prostate cancer risk by tumour grade (χ^2_1=7.35; P_{het}=0.007); low circulating free testosterone concentration was associated with a reduced risk of low-grade prostate cancer (OR=0.76; 95% CI 0.67-0.88), while there was a non-significantly increased risk of high-grade prostate cancer (OR=1.56; 95% CI 0.95-2.57) (Figure 6.5). There was no evidence of heterogeneity in the association by tumour stage, aggressiveness, PSA era, or other characteristics (Figure 6.5).

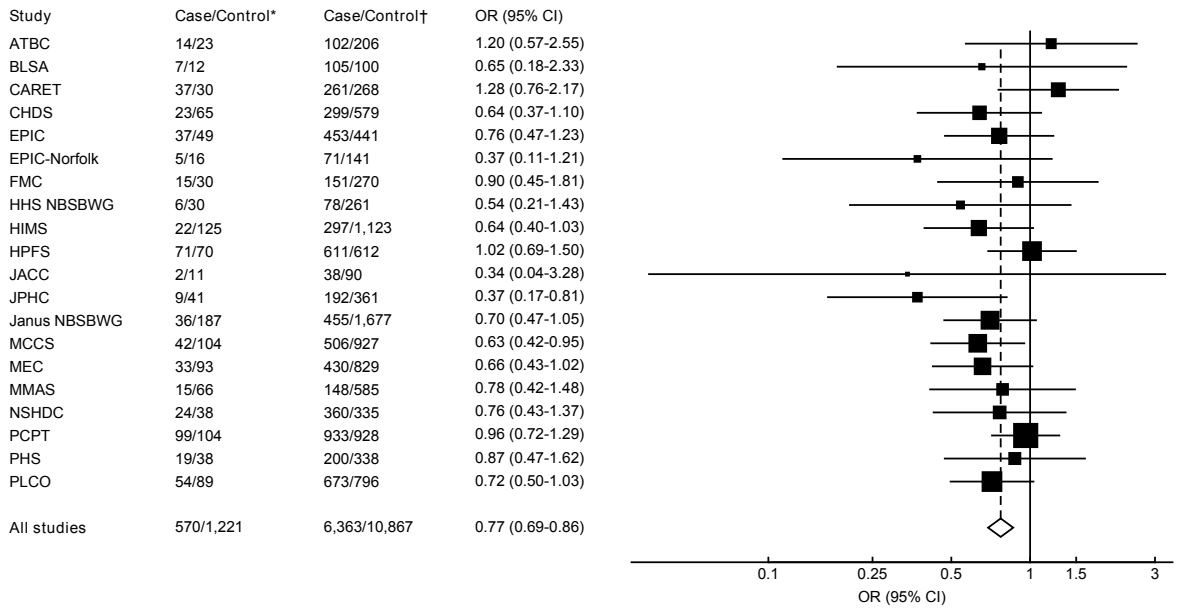


Figure 6.4: Odds ratio (95% CIs) of prostate cancer risk by study-specific free testosterone concentrations (1^{st} vs. 2^{nd}–10^{th}) of study-specific) by study.

Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, and education status.
Test of significance (studies without organised screening): p<0.001.
Test of heterogeneity between studies without organised screening=χ^2_{17}=15.88; p=0.53.
Test of significance (studies with organised screening): p=0.16.
Test of heterogeneity between studies with organised screening=χ^2_1=1.29; p=0.26.
Test of heterogeneity between studies with and without organised screening=χ^2_1=0.73; p=0.39.
Test of significance (overall): p<0.001.
Test of heterogeneity overall=χ^2_{19}=18.0; p=0.53.
*1^{st} study-specific tenth of free testosterone.
†2^{nd}–10^{th} study-specific tenths of free testosterone.
Abbreviations: ATBC=Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BMI=body mass index; BLSA=Baltimore Longitudinal Study of Aging; CARET=Carotene and Retinol Efficacy Trial; CHDS=Child Health and Development Studies; CI=confidence interval; EPIC=European Prospective Investigation into Cancer and Nutrition; FMC=Finnish Mobile Clinic Health Examination Survey; HHS=Helsinki Heart Study; HIMS=Health In Men Study; HPFS=Health Professionals Follow-up Study; JACC=Japan Collaborative Cohort Study; JPHC=Japan Public Health Center-based Prospective Study; MCCS=Melbourne Collaborative Cohort Study; MEC=Multiethnic Cohort Study of Diet and Cancer; MMAS=Massachusetts Male Aging Study; NBSBWG=Nordic Biological Specimen Biobank Working Group; NSHDC=Northern Sweden Health and Disease Cohort; OR=odds ratio; PCPT=Prostate Cancer Prevention Trial; PHS=Physicians' Health Study; PLCO=Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

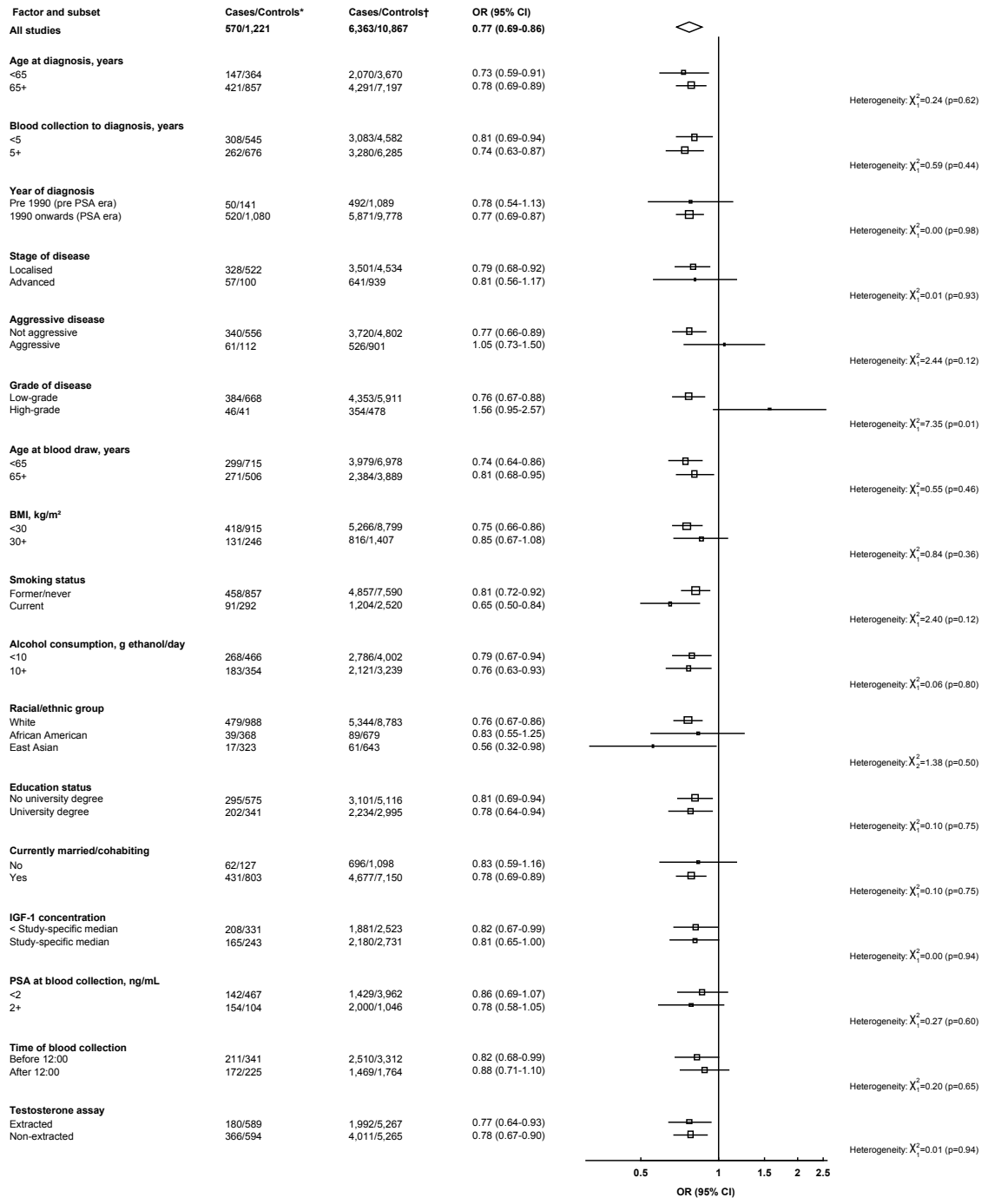


Figure 6.5: Odds ratio (95% CIs) of prostate cancer risk by study-specific tenths of free testosterone concentrations (1^{st} vs. 2^{nd}–10^{th}), by participant characteristics.

Estimates are from logistic regression conditioned on the matching variables and adjusted for age, BMI, height, alcohol intake, smoking status, marital status, and education status.

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 χ^2 test of interaction between subgroup and the binary variable.

*1^{st} study-specific tenth of free testosterone.

†2^{nd}–10^{th} study-specific tenths of free testosterone.

Abbreviations: BMI=body mass index; CI=confidence interval; IGF=insulin-like growth factor; OR=odds ratio; PSA=prostate-specific antigen.

6.3.4 Other assay measures

PSA

Men in the lowest tenth of free testosterone had 22% lower PSA concentrations in comparison with men in the 2^{nd}-10^{th} tenths (Table 6.6 and Figure 6.6.A). PSA concentration at blood collection was available for a subset of participants (prostate cancer cases=3,800). In this subset, men with low free testosterone had a reduced risk of prostate cancer (OR=0.74; 95% CI 0.64-0.85); further adjustment for PSA attenuated the association to the null (OR=0.92; 95% CI 0.76-1.12).

IGF-I

IGF-I has a modest positive correlation with free testosterone (Chapter 4) and men in the lowest study-specific tenth of free testosterone also had lower peripheral IGF-I concentrations (Figure 6.6.B). IGF-I concentrations at blood collection were available for a subset of participants (prostate cancer cases=4,540). In this restricted subset, low free testosterone was associated with prostate cancer risk (OR for 1^{st} versus 2^{nd}-10^{th}=0.76; 95% CI 0.60-0.98), and further adjustment for IGF-I slightly strengthened this association (OR for 1^{st} versus 2^{nd}-10^{th}=0.57; 95% CI 0.37-0.89).

6.3.5 Sensitivity analysis

The association of low free testosterone with prostate cancer was slightly weaker with the use of an alternative method to estimate free testosterone concentration, but the association remained significant (OR for 1^{st} versus 2^{nd}-10^{th} category=0.85, 95% CI 0.76-0.94).

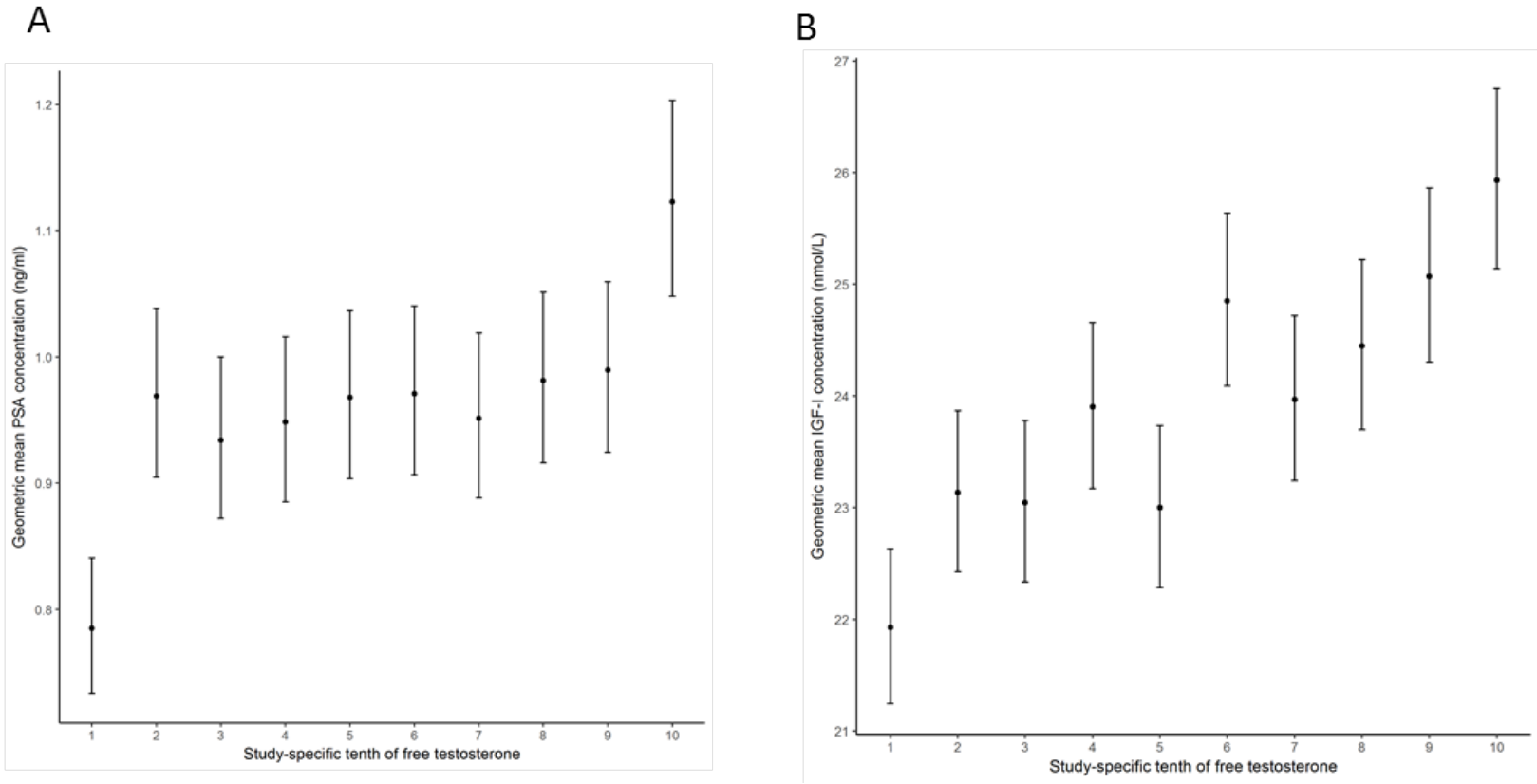


Figure 6.6: Geometric mean hormone concentrations by study-specific tenth of free testosterone, adjusted for age and BMI (controls only).
A: PSA (n=5,730).
B: IGF-I (n=5,933).
Abbreviations: BMI=body mass index; IGF-I=insulin-like growth factor-I; PSA=prostate-specific antigen.

6.4 Discussion

6.4.1 Main findings

These results indicate that men in the lowest study-specific tenth of calculated free testosterone concentration have a 23% reduced risk of prostate cancer compared with men with higher concentrations. Above this very low concentration, prostate cancer risk did not change with higher circulating free testosterone concentrations. There was also evidence that this association varied by tumour grade. This is the largest collection of data on hormones and prostate cancer risk available and is the first large-scale prospective evidence that supports an association between low free testosterone concentrations and prostate cancer risk.

6.4.2 Biological plausibility

The observed association between low free testosterone and lower prostate cancer risk may be due to a direct biological effect. Across the normal range of circulating free testosterone concentrations, the stimulation of the prostatic androgen receptors may remain relatively constant, due to stable intra-prostatic DHT concentrations and/or saturation of androgen receptors (Chapter 2) [19, 196]. Therefore, variation across the normal range of circulating free testosterone concentrations may not be associated with prostate cancer risk. However, when circulating concentrations are very low, reduced androgen receptor signalling may lead to a reduction in prostate cancer risk [19, 196, 431].

6.4.3 Heterogeneity by tumour grade

While there was no evidence of heterogeneity in the association of free testosterone with prostate cancer risk by tumour stage or aggressiveness, there was evidence of heterogeneity in this association by tumour histological grade; low free testosterone concentration was associated with a lower risk of low-intermediate grade prostate cancer, and there was a non-significantly increased risk of high-grade disease. Although it is possible that this heterogeneity is a chance finding due to the multiple tests conducted and the relatively small number of high-grade tumours, this pattern

has been reported previously in the HPFS cohort [245], clinical case studies [432–434], and the PCPT and Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trials [9, 10]. These two randomised controlled trials investigated the effect on prostate cancer risk of 5α-reductase inhibitors, which can reduce intra-prostatic DHT concentrations by approximately 80-90% [427]. The PCPT trial reported a 25% reduction in overall prostate cancer risk with use of finasteride, but this was also associated with a 27% increase in the incidence of high-grade (Gleason grade $\geqslant$7) tumours, based on 517 high-grade cancers (P=0.005) [10]. The REDUCE trial reported an overall 23% reduction in risk for prostate cancer with the use of dutasteride, but a 58% increased risk of high-grade (Gleason grade $\geqslant$8) tumours, although the number of such tumours was low (n=48; 3.2% of all cases) [9].

There are several possible explanations for the observed heterogeneity in the associations by tumour grade. Prostate tumour grade stays stable over several years [435], suggesting that high-grade tumours may develop *de novo* rather than from the de-differentiation of low-intermediate grade tumours. Mechanistically, prostatic androgen-androgen receptor binding is an important modulator of cell differentiation [183], thus prostate cells with a lower androgen exposure may be less differentiated and more likely to develop into high-grade tumours [436]. Alternatively, the observed heterogeneity may be a differential growth adaptation of early low-grade cancer lesions to a low androgen environment.

Another possibility is differential detection bias, due to a lower prostatic volume (and hence increased tumour sampling density), changes in tumour morphology, or sensitivity to PSA testing, as discussed in relation to PCPT and REDUCE [437–444]. Owing to the clinical importance of high-grade tumours, this observed heterogeneity by grade, with a possible higher risk of high-grade tumours, requires further investigation.

6.4.4 Detection bias

The association with overall prostate cancer may be partially explained by detection bias. Androgen receptor signalling directly stimulates PSA expression [194, 204, 445]. Therefore, large reductions in circulating testosterone concentrations may reduce serum PSA [191]. In this analysis controls in the lowest tenth of free testosterone had lower PSA concentrations at blood collection. This observation is consistent with the hypothesis that men with low circulating free testosterone have lower androgen receptor signalling and supports the assumption in this analysis that men in this tenth have sufficiently low circulating free testosterone concentrations to reduce intra-prostatic androgen signalling. Adjustment for PSA concentration in a subset of participants attenuated the association of low free testosterone and prostate cancer risk towards the null. This may indicate that men with low free testosterone have lower PSA concentrations and thus are less likely to have follow-up tests, therefore prostate tumours are less likely to be detected. This may also explain the possible heterogeneity by tumour grading, as only more clinically relevant tumours are identified leading to differential detection bias. However, there was no evidence of heterogeneity in the associations between men diagnosed pre and post 1990 (before 1990 there was relatively little PSA testing) [446]. The direct regulation of PSA by androgen receptor signalling, its use in diagnosing prostate cancer and its increasing serum PSA concentrations with prostate tumour development and progression means that it is not possible to disentangle the relationship between these variables in this observational analysis.

6.4.5 Co-morbidities

Co-morbidities may play a role in the association between free testosterone and prostate cancer, particularly type II diabetes and/or obesity. Both these diseases are associated with hypogonadism [447–449] and these men have a lower risk of overall prostate cancer [74, 92, 110, 450] and lower PSA concentrations [113, 451]. Obese men may also have an increased risk of high-grade prostate cancer [92, 450]. The results presented in this analysis are adjusted for BMI, and there was no evidence

of heterogeneity in the associations of free testosterone with prostate cancer risk by whether or not participants were obese. Diabetes status data in the EHNBPCCG were not available at the time of this analysis. The lower circulating concentrations of free testosterone in these diabetic men may be a potential mechanism which explains the observed reduction in prostate cancer risk in men with these co-morbidities, or may affect how likely an individual is to attend prostate testing or biopsy, and may therefore contribute to the differential detection tumour grade [452].

6.4.6 IGF-I at blood collection

Higher circulating IGF-I concentrations are associated with an elevated prostate cancer risk [14] and is correlated with free testosterone concentrations (Chapter 5). Therefore, it is possible that the association with free and not total testosterone may be accounted for by lower IGF-I concentrations in these men. However, there was no evidence of heterogeneity in the associations of free testosterone with prostate cancer by IGF-I concentrations at blood collection, and further adjustment for IGF-I slightly strengthened the association of free testosterone with prostate cancer risk.

6.4.7 Free hormone hypothesis

This analysis assumes that only free testosterone, which is not bound to either albumin or SHBG, is able to exert an androgenic effect [158, 159]. This is known as the free hormone hypothesis (Chapter 2).

Free testosterone was calculated using the law of mass action [163, 302], which is based on testosterone and SHBG concentrations and assumes a constant albumin concentration. Although this is a commonly used method of estimating free testosterone concentrations, it has not been validated within each individual study via equilibrium dialysis, which is considered the reference method [159]. Other studies have sought to validate this calculation against equilibrium dialysis, finding good agreement [350]. However, calculations are dependent on assumptions of binding dynamics, stoichiometry and affinity constants [159], which may be biased by factors such as hormone treatment and adiposity [350]. Other estimations of free

testosterone have used statistical models to predict free testosterone concentrations and validated this against a measured sample [306, 347]. These empirical calculations are likely to be be specific to their study population and laboratory methodology, whilst the calculations based on the law of mass action may be more generalisable. Use of the Sartorius empirical calculation as an alternative method to estimate of free testosterone concentrations [306, 350] did not materially affect the associations. This is probably attributable to the classification of participants based upon their relative distribution as opposed to absolute clinical cut-points. Therefore, men with low testosterone and high SHBG concentrations are likely to be classified as having low free testosterone regardless of the method of estimation.

6.4.8 Strengths and limitations

The primary strength of this analysis is its large sample size, which increased the statistical power to detect associations at more extreme hormone concentrations. This individual participant analysis enables the detection of associations which would not be possible via a standard meta-analysis and is based on both published and unpublished data which reduces the possibility of publication bias.

This study has some limitations. The assay methods used to measure testosterone and SHBG varied, with most studies using non-extraction assays to measure testosterone. While this may introduce some misclassification, this would be expected to be non-differential and therefore tend to bias any association towards the null. Mass spectrometry is often considered the gold standard method to measure sex hormone concentrations [344], but high quality immunoassays are able to reliably measure low adult male testosterone concentrations [429, 453]. The HIMS was the only included study which used mass spectrometry to measure serum testosterone, and the association was not materially different to studies which used immunoassays. As with the free testosterone calculation, these assays may not be suitable for determining absolute clinical cut-points, but they may be appropriate to determine relative concentrations within studies [429].

Variation in individual study designs and data collection both at baseline, participant follow-up and disease characterisation, has meant that there is a relatively large amount of missing data, which limited the ability to adjust for potential confounders. However, only a small number of prostate cancer risk factors have been identified (Chapter 3), suggesting that the likelihood of substantial confounding is low. Moreover, the analyses in Chapter 4 showed only modest associations of free testosterone with factors other than age and BMI, again suggesting that the potential for residual confounding may be small. The number of high-grade and/or aggressive tumours in these analyses was also limited.

There was variation in biomarker measurements between studies due to differences in blood collection, storage and assays [428]. Therefore, these analyses categorised hormone concentrations into study-specific tenths, which assumes that the true range of hormone concentrations within each study is similar. Whilst this method is more easily interpretable, it is not possible to quantify the possible threshold risk concentration. Studies which have examined the effects of testosterone inhibition on intra-prostatic androgens and/or prostate function have generally focused on the effects of medical castration [191]. The minimal free testosterone concentrations required to significantly reduce intra-prostatic DHT concentration are not known. Due to a lack of statistical power it was not possible to group the population into smaller categories.

This study also relied on single measurements of testosterone and SHBG, with an average time from blood collection to diagnosis of 6.8 years, to represent participants' hormone concentrations over the medium to long term. While several studies show that a single measure of these analytes has moderately good reproducibility over a period of up to five years [355–358] (e.g. for testosterone r=0.85 over a one year period) [355], it is unknown whether these measures are reliable over the longer term. Random measurement error may attenuate the associations due to regression dilution bias [290].

6.5 Conclusion

In summary, the findings from this pooled prospective analysis of 6,933 prostate cancer cases and 12,088 controls support the hypothesis that very low concentrations of circulating free testosterone are associated with a reduced risk of overall prostate cancer, but these men may have an increased risk of high-grade prostate cancer. Further research is needed to elucidate whether these associations are causal or due to detection bias and explore the apparent differential association by tumour grade.

7

Hormone-related diseases and prostate cancer risk

Contents

7.1 Introduction 180
7.1.1 Hormone-related disorders 180
7.1.2 Aims 183
7.2 Methods 184
7.2.1 Study design and population 184
7.2.2 Statistical analysis 189
7.2.3 Statistical software 190
7.3 Results 190
7.3.1 Reference cohort 190
7.3.2 Acromegaly cohort 190
7.3.3 Hypogonadal cohort 193
7.3.4 Further analyses 193
7.4 Discussion 198
7.4.1 Acromegaly cohort 198
7.4.2 Hypogonadal cohort 198
7.4.3 Detection bias 200
7.4.4 Strengths and limitations 201
7.5 Conclusion 203

7.1 Introduction

The analysis in Chapter 6 using the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG) dataset, found that men with low free testosterone may have a lower risk of overall prostate cancer. Previous pooled analyses by the EHNBPCCG have also indicated that circulating insulin-like growth factor-I (IGF-I) concentrations have positive linear associations with prostate cancer risk [14] (Chapter 2).

Further analyses are required to test whether these observed associations with circulating hormone concentrations are robust across different contexts. Prospective cohort studies are generally underpowered to investigate extremes of the distributions and absolute clinical cut-points are difficult to define using heterogeneous data from international consortia. Hospital records are routinely collected across all English National Health Service (NHS) hospital admissions and can also be linked to other national databases including mortality records (Chapter 3). Between April 2016 and March 2017 16.5, million hospital admissions were recorded (Chapter 3) [454]. Using this nationally collected data it is possible to identify men who have been diagnosed with hormone-related diseases and "follow-up" these men to estimate the subsequent risk of prostate cancer diagnosis and mortality.

Analyses of prostate cancer risk in men with hormone-related diseases will provide further evidence on the hypothesised of these hormones as risk factors for prostate cancer and therefore, will contribute to the understanding of the possible roles of IGF-I and testosterone in prostate cancer development.

7.1.1 Hormone-related disorders

Several diseases are characterised by abnormal circulating hormone concentrations. Acromegaly is characterised by high IGF-I concentrations, while Klinefelter's syndrome, testicular hypofunction and hypopituitarism are associated with clinically low testosterone concentrations (hypogonadism).

High IGF-I disorders

Acromegaly: is characterised by the hypersecretion of growth hormone by the pituitary gland, generally triggered by a pituitary tumour. This increase in growth hormone secretion leads to increased IGF-I production. Symptoms include acral overgrowth and swelling of the internal organs. The prevalence of acromegaly is approximately 130 per million population [455].

Acromegaly is diagnosed by the failure to suppress growth hormone during an oral glucose tolerance test and elevated circulating IGF-I concentrations [456]. A clinical study of 97 patients diagnosed with acromegaly reported that circulating IGF-I concentrations were on average 2.31 $\pm$ 1.1 times greater than the upper limit of normal after accounting for age and sex [457]. Long-term biochemical control of the condition is estimated to be achieved in 50-65% of patients who receive either surgical or hormone treatment [458].

Acromegaly may be associated with an elevated risk of overall cancer, but associations with prostate cancer are unclear [459].

Low testosterone disorders

There are three clinically defined diagnostic codes that describe diseases which are characterised by hypogonadism.

Klinefelter's syndrome: a genetic abnormality due to two or more X chromosomes in males. It is the most common genetic cause of human male infertility and its prevalence is estimated to be 1 in 660 men [460].

The mechanism by which this aneuploidy leads to hypogonadism is poorly understood. Symptoms include gonadal dysfunction and gynecomastia [461, 462]. It is estimated that between 50-75% of Klinefelter's syndrome cases are not diagnosed [461, 463, 464], possibly due to less serve phenotypic presentations. Suspected diagnosis can be confirmed by chromosomal analysis [465].

Circulating total testosterone concentrations are low in 45-85% of men diagnosed with Klinefelter's syndrome [460, 461, 466, 467]. The normal range of circulating total testosterone concentrations in men is 12-40 nmol/L and free testosterone is

>250 pmol/L [468]. A clinical study of men diagnosed with Klinefelter's syndrome reported that these men had median circulating total testosterone concentration=9.9 nmol/L (interquartile range (IQR)=7.0-16.2 nmol/L) and free testosterone median concentration=205.0 pmol/L (IQR=127.5-279.5 pmol/L) (n=126-186) [465].

Observational studies have reported that men diagnosed with Klinefelter's syndrome may have a lower risk of being diagnosed with prostate cancer [463, 469, 470] and may also have a lower risk of prostate cancer mortality [463].

Testicular hypofunction: describes impaired testicular function which leads to reductions in testosterone production (also known as primary hypogonadism). Testicular hypofunction encompasses a broad range of disorders which can be caused by genetic mutations (including Klinefelter's syndrome) and environmental exposures including illness and haemochromatosis [471–473]. Clinical diagnosis is made on the basis of signs and symptoms which are consistent with androgen deficiency as well as low measured testosterone concentrations (typically below 8-10 nmol/L) and elevated gonadotrophin concentrations [468]. In European men aged 40-79 years, the prevalence of primary hypogonadism is estimated to be ∼2% [474].

A clinical study of men with primary hypogonadism (n=63), who were not receiving hormone medication, reported mean circulating total testosterone concentrations=7.5 nmol/L (standard deviation (SD)=2.5 nmol/L) and mean free testosterone=138.7 pmol/L (SD=59.8 pmol/L) in these men [474].

It is currently unknown whether diagnosis of testicular hypofunction is related to prostate cancer risk.

Hypopituitarism: is characterised by the complete or partial insufficiency of anterior pituitary hormone secretions. Prevalence of hypopituitarism is estimated to be ∼29-45.5 per 100,000 [475]. This condition has a number of different causes which can lead to onset in childhood (genetic disorders, developmental abnormalities and perinatal damage) and adulthood (pituitary tumours, irradiation and trauma) [476, 477]. Symptoms vary due the degree of the hormone deficiency and include growth retardation and infertility.

The anterior pituitary gland secretes the gonadotropic hormones (follicle-stimulating hormone (FSH) and luteinising hormone (LH)) as well as other hormones (including adrenocorticotropic hormone and somatostatin). A study of adult Caucasian men reported 87% of men diagnosed with hypopituitarism were FSH and/or LH deficient [475]. Low concentrations of these gonadotropins lead to a decline in testosterone secretion via the hypothalamic-pituitary-gonadal axis (Chapter 2). Hypogonadism related to the hypothalamus or the pituitary gland is known as secondary hypogonadism. Because hypopituitarism encapsulates a broad range of diseases it is difficult to estimate testosterone deficiency specifically in these men. A study of men with secondary hypogonadism (n=375) reported that these men had mean total testosterone=8.7 nmol/L (SD=1.6 nmol) and mean free testosterone=194.4 pmol/L (SD=45.6 pmol/L) [474].

It is currently unknown whether diagnosis of hypopituitarism is related to prostate cancer risk.

7.1.2 Aims

This analysis aimed to test whether there was an association between i) high IGF-I (acromegaly) and ii) low testosterone (Klinefelter's syndrome, testicular hypofunction and hypopituitarism) with subsequent risk of prostate cancer diagnosis and mortality.

7.2 Methods

7.2.1 Study design and population

English national Hospital Episode Statistics Admitted Patient Care (HES APC) and mortality data from 1st January 1999 to 31st March 2017 were used to construct and follow-up several cohorts of patients through record linkage. HES APC data comprise demographic and clinical data recorded for all patients undergoing admission to NHS hospitals and treatment centres in England. A hospital admission includes any visit to a hospital which requires a hospital bed. This includes emergency and planned admission and day cases [267], but excludes accident and emergency attendance (A & E) or outpatient bookings. These records are linked so that individuals can be traced through multiple admissions over time. Further details of HES APC records and data linkage can be found in Chapter 3.

Acromegaly and hypogonadal cohorts

The "exposure" cohorts were constructed to define the population of males diagnosed with acromegaly, Klinefelter's syndrome, testicular hypofunction and hypopituitarism. Cohorts of men were identified in the linked HES APC dataset using the International Classification of Diseases, 10th revision (ICD-10) codes E22.0 (acromegaly), Q98.0-Q98.4 (Klinefelter's syndrome), E29.1 (testicular hypofunction), and E23.0 (hypopituitarism). Klinefelter's syndrome, testicular hypofunction and hypopituitarism are all conditions that are characterised by low testosterone concentrations; therefore, these conditions were also combined into one cohort to maximise statistical power.

The inclusion criteria were men aged 35 years or more who had an admission during the study period with one of the exposure conditions, coded in any diagnosis position on the discharge record (i.e. either as a primary diagnosis or as a co-morbidity during an admission for another medical or surgical procedure). Date of entry was based on the patient's earliest known date of hospital admission for the relevant condition.

All participants who had a prior prostate cancer diagnosis, who had both acromegaly and hypogonadal-associated diseases on their hospital records, or had missing covariate data (age, year of cohort entry, region of residence, Index of Multiple Deprivation (IMD) rank) were excluded.

Reference cohort

A reference cohort of men with no known prior record of prostate cancer was constructed by identifying men aged 35 years or more without acromegaly, Klinefelter's syndrome, testicular hypofunction or hypopituitarism, who were admitted to hospital for various conditions and injuries recorded as the primary diagnosis or main operation on the hospital record: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement or tonsillectomy. This diverse range of conditions was chosen on the basis that they are common and relatively minor, so that the men in the reference cohort would be broadly representative of the general population. Conditions/operations were recorded as the primary diagnosis or operation in the hospital record to avoid selecting men who were coming into hospital principally for more serious or uncommon problems.

Date of entry to the reference cohort was based on the patient's earliest known date of admission for one of these conditions.

Follow-up for prostate cancer diagnosis and vital status

From cohort entry, men were followed-up through record linkage for any subsequent day-case or inpatient admission for prostate cancer (ICD-9 code 185, ICD-10 C61) [478], recorded either as a primary diagnosis or elsewhere on the hospital record. To censor for death and to identify further prostate cancer diagnoses not otherwise recorded in hospital, the HES APC records were linked to mortality records obtained from the Office for National Statistics' Register of Deaths in England and Wales (further details in Chapter 3), and assigned prostate cancer case if it appeared

anywhere on the death record (Table 7.1). This register was also used to identify prostate cancer mortality. Tumour subtype data were not available in these datasets, therefore prostate cancer mortality was used as a marker of tumour aggressiveness. To capture these clinically aggressive tumours, prostate cancer death was defined as the underlying cause of death only (Table 7.1).

This analysis was not linked to the cancer registry data. To estimate the extent of under-ascertainment (i.e. the proportion of prostate cancer patients who were not identified using HES APC records or mortality data), the proportion of men with prostate cancer recorded in cancer registry data that did not have a concurrent prostate cancer HES APC and/or mortality record was calculated in the UK Biobank dataset. In this separate dataset HES APC and death data combined captured 76% of prostate cancer cases observed in the English cancer registry.

Table 7.1: Identification of prostate cancer by HES APC and death records.

Endpoint		High IGF-I cohort	Hypogonadal cohort			Reference cohort
		Acromegaly	Klinefelter's syndrome	Testicular hypofunction	Hypopituitarism	Reference cohort conditions†
Prostate cancer cases	Total	96	19	140	227	127,334
	Death record only*	4 (4.2%)	0 (0.0%)	1 (0.7%)	8 (3.5%)	5,737 (4.5%)
Prostate cancer deaths	Primary cause of death only	19	2	13	23	29,022

*Anywhere on death record.
% value denotes the percentage of the total cases that were identified from death record only (i.e. no hospital record of prostate cancer diagnosis).
† Conditions used in the reference cohort: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement, tonsillectomy.
Abbreviations: HES APC=Hospital Episode Statistics Admitted Patient Care; IGF-I=insulin-like growth factor-I

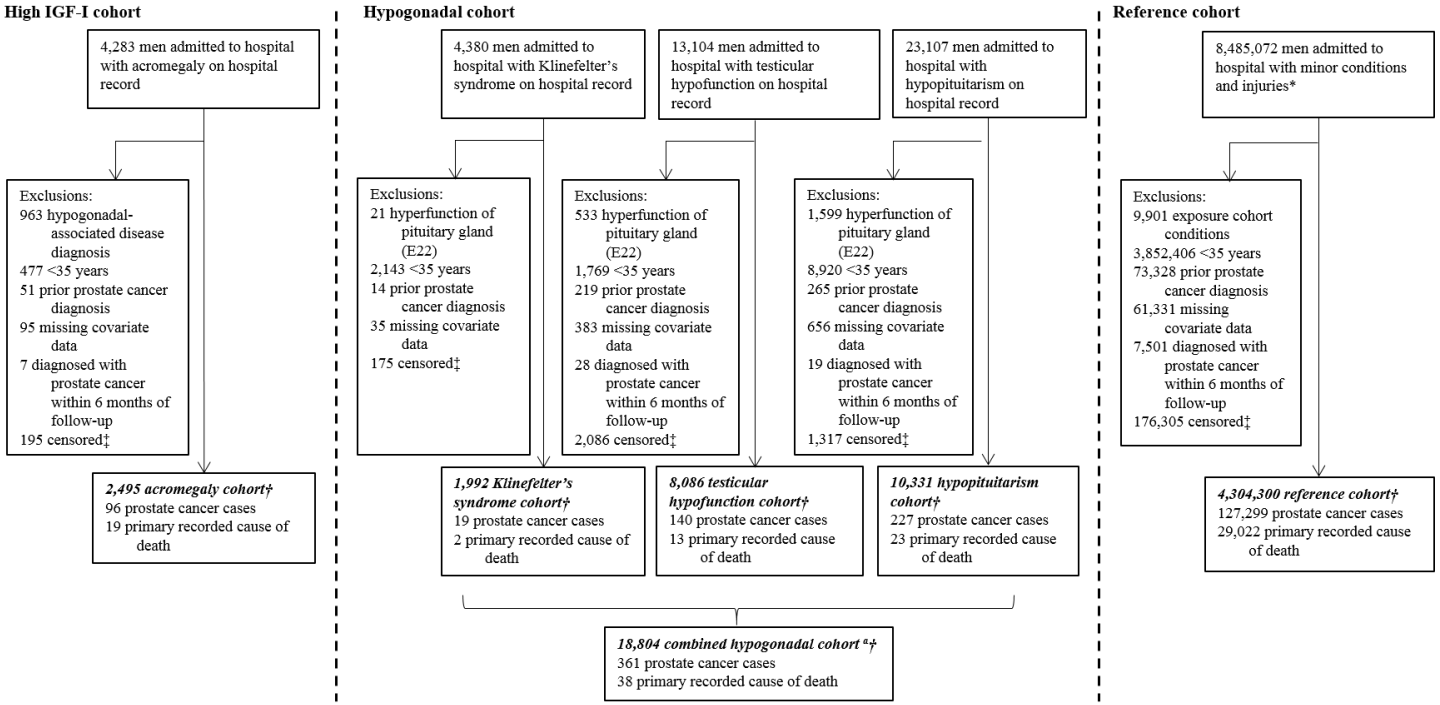


Figure 7.1: Schematic of cohort selection criteria.

* Conditions used in the reference cohort: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement, tonsillectomy.

† Cohort numbers refer to the men included in the prostate cancer incident analysis. For the analysis of prostate cancer mortality men who were diagnosed with prostate cancer within the first 6 months of follow-up but did not die within this time were included in this analysis. These men were excluded in the analysis of incident prostate cancer.

‡ Censoring due to men who died within the first 6 months of follow-up, or those who reached the end of study follow-up (31st March 2017) within the first 6 months of follow-up.

[a] Hypogonadal disease cohorts are not mutually exclusive therefore the combined hypogonadal cohort may sum to less than the individual hypogonadal disease cohorts.

7.2.2 Statistical analysis

For each patient the baseline date for cohort entry was their earliest known admission date for acromegaly, a hypogonadal condition or a reference cohort disease (from 1st January 1999 onwards). Eligible men contributed person-years from baseline until the date of first known prostate cancer diagnosis, death, or the end of the follow-up period (31st March 2017), whichever came first.

Multivariable Cox proportional hazards regression models were used to estimate hazard ratios (HRs) of i) incident prostate cancer and ii) prostate cancer mortality; comparing in turn each exposed cohort with the reference cohort. HRs were adjusted for age (in five year groups), calendar year of first recorded admission (to remove period effects), region of residence (nine regions), and patients' IMD (in quintiles) [479]. Time from cohort entry was used as the underlying time variable.

In an initial analysis, examination of Schoenfeld residuals indicated that the proportional hazards assumption was violated, due to an increase in prostate cancer diagnoses within the first 6 months of cohort entry. Therefore, all participants who were diagnosed with prostate cancer within the first 6 months of cohort entry across all cohorts were excluded. After excluding these men, the analytical cohorts included 2,495 men with acromegaly, and 1,992, 8,086, and 10,331 men with Klinefelter's syndrome, testicular hypofunction, and hypopituitarism, respectively, and 4.3 million men in the reference cohort (Figure 7.1).

Sub-group analyses were also conducted to assess whether associations differed by age at cohort entry (35-64; 65+ years) and by length of follow-up (6 months-4 years; 5+ years), these categories were chosen *a priori* due to possible differences in prostate tumours in younger men [480] and detection bias [481]. For these purposes, interaction terms were fitted between the subgroup variables and prostate cancer diagnosis, and the models were compared with and without the interaction terms using likelihood ratio (LR) tests to determine statistical significance.

7.2.3 Statistical software

All statistical tests were two-sided, with the statistical significance threshold set at $p<0.05$. Data analyses used Stata Statistical Software release 14.1 (Stata Corp., College Station, TX, USA), and figures were created in R version 3.2.3.

7.3 Results

Participant characteristics at cohort entry are displayed in Table 7.2. Most men entered the cohort aged 35-64 years. A higher proportion of men diagnosed with acromegaly entered the cohort Jan 1999-2003 than the reference cohort, whereas the proportion of men diagnosed with testicular hypofunction and hypopituitarism increased with time. Men in the Klinefelter's syndrome cohort were younger and had higher levels of socioeconomic deprivation in comparison with the other cohorts (Table 7.2).

7.3.1 Reference cohort

In the reference cohort, median follow-up time was 7.7 years (IQR 3.9-12.4), and included approximately 4.3 million men, of whom 127,299 were diagnosed with prostate cancer and 29,022 died from the disease (as the underlying cause of death).

7.3.2 Acromegaly cohort

In the acromegaly cohort, median follow-up time was 7.7 years (IQR 3.5-13.3), during which time 96 cases of prostate cancer were diagnosed and 19 men died from prostate cancer. Compared with the reference cohort, acromegaly was associated with a 33% increased risk of being diagnosed with prostate cancer (HR=1.33, 95% CI 1.09-1.63; P=0.005), and a 44% (non-significant) increased risk of prostate cancer death (HR=1.44, 95% CI 0.92-2.26; P=0.11) (Figure 7.2).

Table 7.2: Participant characteristics at cohort entry*.

	Acromegaly (n=2,495)	Klinefelter's Syndrome (n=1,992)	Testicular hypofunction (n=8,086)	Hypopituitarism (n=10,331)	Reference cohort (n=4,304,300)
Age group years, n (%)					
35-49	698 (28.0)	916 (46.0)	2,486 (30.7)	2,771 (26.8)	1,441,418 (33.5)
50-64	1,021 (40.9)	691 (34.7)	3,421 (42.3)	3,782 (36.6)	1,303,855 (30.3)
65-79	675 (27.1)	349 (17.5)	1,883 (23.3)	3,009 (29.1)	1,184,098 (27.5)
80+	101 (4.1)	36 (1.8)	296 (3.7)	769 (7.4)	374,929 (8.7)
Year of index admission, n (%)					
Jan 1999-2003	1,007 (40.4)	513 (25.8)	1,213 (15.0)	2,062 (20.0)	1,265,482 (29.4)
2004-2008	590 (23.7)	472 (23.7)	1,563 (19.3)	2,265 (21.9)	1,208,125 (28.1)
2009-2013	584 (23.4)	617 (31.0)	3,013 (37.3)	3,581 (34.7)	1,202,353 (27.9)
2014- March 2017	314 (12.6)	390 (19.6)	2,297 (28.4)	2,423 (23.5)	628,340 (14.6)
Government office region, n (%)					
North East	104 (4.2)	167 (8.4)	509 (6.3)	633 (6.1)	264,976 (6.2)
North West	305 (12.2)	326 (16.4)	1,709 (21.1)	1,774 (17.2)	643,095 (14.9)
Yorkshire and Humber	258 (10.3)	203 (10.2)	1,146 (14.2)	1,222 (11.8)	452,720 (10.5)
East Midlands	177 (7.1)	145 (7.3)	437 (5.4)	693 (6.7)	357,496 (8.3)
West Midlands	212 (8.5)	172 (8.6)	717 (8.9)	723 (7.0)	441,696 (10.3)
East of England	277 (11.1)	236 (11.9)	754 (9.3)	1,023 (9.9)	466,580 (10.8)
London	404 (16.2)	229 (11.5)	987 (12.2)	1,648 (16.0)	522,824 (12.2)
South East	464 (18.6)	294 (14.8)	1,131 (14.0)	1,533 (14.8)	658,130 (15.3)
South West	294 (11.8)	220 (11.0)	696 (8.6)	1,082 (10.5)	496,783 (11.5)
Index of multiple deprivation quintile, n (%)					
1 (highest deprivation)	463 (18.6)	523 (26.3)	1,831 (22.6)	2,112 (20.4)	796,636 (18.5)
2	461 (18.5)	451 (22.6)	1,632 (20.2)	2,042 (19.8)	832,001 (19.3)
3	535 (21.4)	395 (19.8)	1,567 (19.4)	2,080 (20.1)	897,729 (20.9)
4	519 (20.8)	371 (18.6)	1,624 (20.1)	2,094 (20.3)	916,192 (21.3)
5 (lowest deprivation)	517 (20.7)	252 (12.7)	1,432 (17.7)	2,003 (19.4)	861,742 (20.0)

*5 cohorts in total are used in this analysis. All data excludes men diagnosed with prostate cancer within the first 6 months of follow-up.

Hormone-associated diseases	Prostate cancer cases/ exposed cohort	Prostate cancer cases/ reference cohort	HR (95% CI)	p-value
IGF-I-associated diseases				
Acromegaly				
Incidence	96/2,495	127,299/4,304,300	1.33 (1.09, 1.63)	0.005
Death	19/2,502	29,022/4,311,029	1.44 (0.92, 2.26)	0.11
Hypogonadal-associated diseases				
All hypogonadal diseases				
Incidence	361/18,763	127,299/4,304,300	0.94 (0.85, 1.05)	0.27
Death	38/18,804	29,022/4,311,029	0.57 (0.42, 0.79)	0.001
Klinefelter's syndrome				
Incidence	19/1,992	127,299/4,304,300	0.58 (0.37, 0.91)	0.02
Death	Insufficient data			
Testicular hypofunction				
Incidence	140/8,086	127,299/4,304,300	0.97 (0.82, 1.15)	0.73
Death	13/8,112	29,022/4,311,029	0.62 (0.36, 1.06)	0.08
Hypopituitarism				
Incidence	227/10,331	127,299/4,304,300	0.98 (0.86, 1.12)	0.81
Death	23/10,347	29,022/4,311,029	0.53 (0.35, 0.79)	0.002

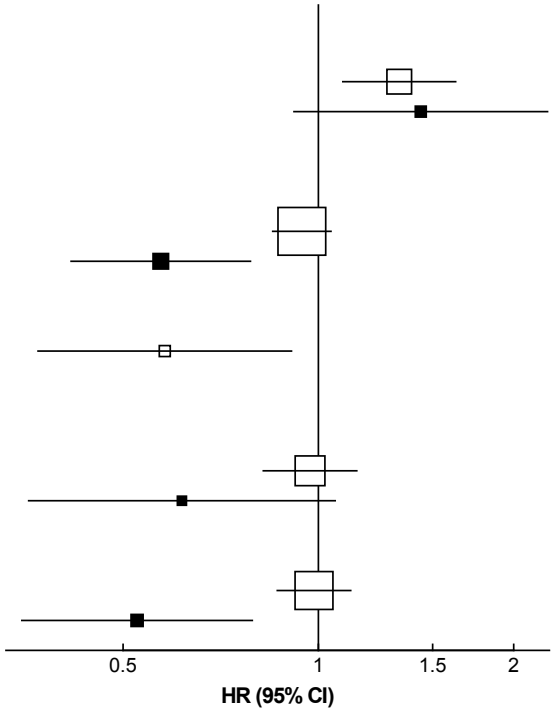


Figure 7.2: Hazard ratio* of prostate cancer diagnosis and mortality in men diagnosed with acromegaly, Klinefelter's syndrome, testicular hypofunction and hypopituitarism, in comparison with the reference cohort†.

*HRs adjusted for age (five year groups), year of cohort entry, region of residence (nine regions), IMD rank (quintiles). These results exclude men diagnosed with prostate cancer within the first 6 months of follow-up.

†Conditions used in the reference cohort: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement, and tonsillectomy.

Abbreviations: CI=confidence interval; HR=hazard ratio; IGF-I=insulin-like growth factor; IMD=index of multiple deprivation.

7.3.3 Hypogonadal cohort

In the hypogonadal cohorts, median follow-up times were 6.8 (IQR 3.2-11.6), 4.9 (IQR 2.3-8.9) and 5.1 years (IQR 2.4-9.2) for the Klinefelter's syndrome, testicular hypofunction and hypopituitarism cohorts, respectively. During these times 19, 140 and 227 men, respectively were diagnosed with prostate cancer, and 2, 13, and 23 prostate cancer deaths were recorded in each cohort, respectively. When all three hypogonadal cohorts were combined, there was no evidence of an association with risk of prostate cancer diagnosis (HR=0.94; 95% CI 0.85-1.05; P=0.27; based on 361 incident cases), but there was a significant reduction in prostate cancer death (HR=0.57, 95% CI 0.42-0.79; P=0.001; based on 38 deaths).

Compared with the reference cohort, men diagnosed with Klinefelter's syndrome had a 42% lower risk of being diagnosed with prostate cancer (HR=0.58, 95% CI 0.37-0.91; P=0.02) (Figure 7.2). There was an insufficient number of prostate cancer deaths to investigate the association of Klinefelter's syndrome with prostate cancer mortality. Testicular hypofunction was not associated with prostate cancer diagnosis (HR=0.97, 95% CI 0.82-1.15; P=0.73), but there was a suggestion of a reduction in prostate cancer mortality, based on 13 deaths (HR=0.62, 95% CI 0.36-1.06; P=0.08) (Figure 7.2). Similarly, hypopituitarism was not associated with overall prostate cancer diagnosis (HR=0.98, 95% CI 0.86-1.12; P=0.81), but was associated with a relative reduction in prostate cancer mortality, based on 23 deaths (HR=0.53, 95% CI 0.35-0.79; P=0.002) (Figure 7.2).

7.3.4 Further analyses

There was no evidence of heterogeneity in the associations of acromegaly or Klinefelter's syndrome with prostate cancer diagnosis by age at cohort entry or length of follow-up (Figures 7.3 and 7.4). There was evidence of heterogeneity in the associations of testicular hypofunction and hypopituitarism with prostate cancer diagnosis by age at cohort entry (χ^2 test for heterogeneity P=0.02 and 0.003, respectively; Figure 7.4). In both cohorts, only men with testicular hypofunction and hypopituitarism who were aged 65+ years at cohort entry had a reduced risk

of prostate cancer diagnosis in comparison with the reference cohort (Figure 7.4). Men aged 35-65 years at cohort entry who were diagnosed with hypopituitarism had a significantly increased risk of being diagnosed with prostate cancer (HR=1.26, 95% CI 1.05-1.52). There was also some indication that men aged 35-64 years who were diagnosed with testicular hypofunction also had an increased risk of prostate cancer diagnosis, but this was not statistically significant (HR=1.19, 95% CI 0.97-1.47) (Figure 7.4).

In the hypopituitarism cohort, there was evidence of heterogeneity in the associations with prostate cancer diagnosis by length of follow-up, with a higher risk of prostate cancer diagnosis within 4 years of diagnosis of hypopituitarism (HR=1.14, 95% CI 0.96-1.52) and a lower risk thereafter (HR=0.83, 95% CI 0.68-1.01; χ^2 test for heterogeneity P=0.03) (Figure 7.4).

Results including men who were identified as having prostate cancer within the first 6 months of cohort entry are displayed in Figure 7.5. Men diagnosed with testicular hypofunction had a 2.5 fold elevated risk of being diagnosed with prostate cancer within six months of cohort entry (95% CI 1.73-3.64). There was no evidence of an increased risk of prostate cancer diagnosis within 6 months of cohort entry for men diagnosed with Klinefelter's syndrome or hypopituitarism (Figure 7.5).

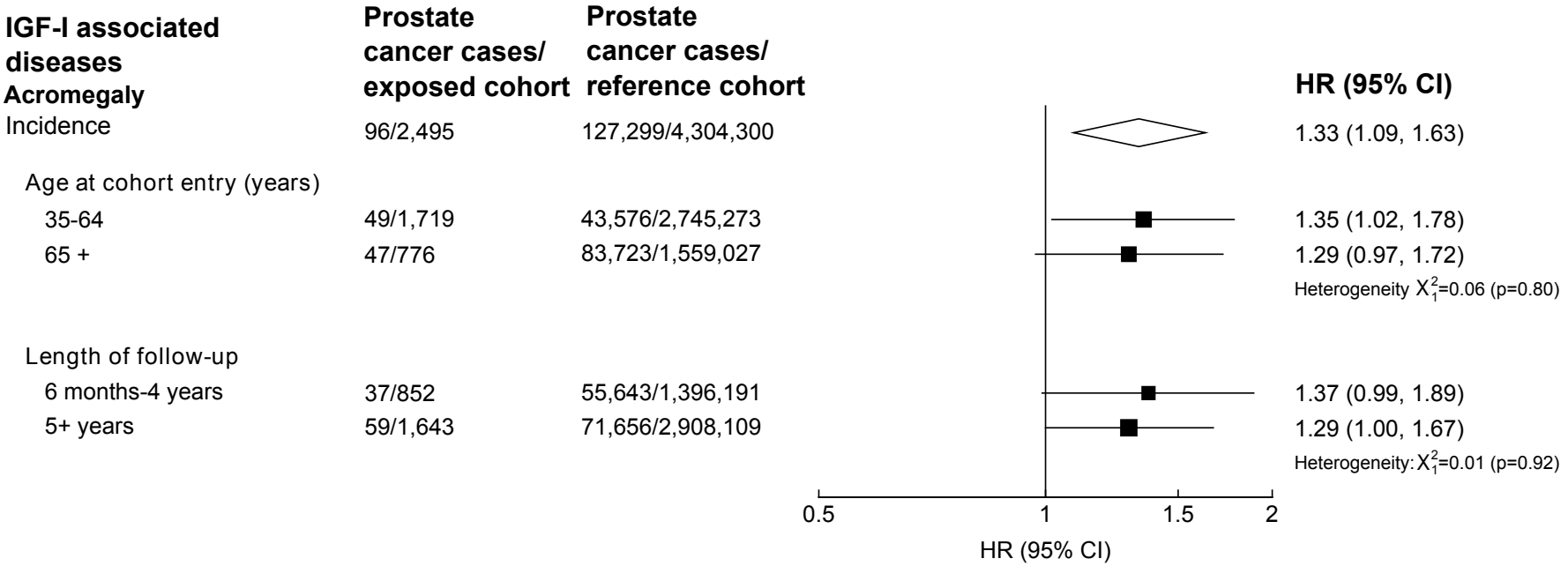


Figure 7.3: Hazard ratio* of prostate cancer diagnosis in men diagnosed with acromegaly in comparison with the reference cohort†, stratified by age at cohort entry and time interval.

*HRs adjusted for age (five year groups), year of cohort entry, region of residence (nine regions), IMD rank (quintiles). Results exclude men diagnosed with prostate cancer within the first 6 months of follow-up.

† Conditions used in the reference cohort: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement, and tonsillectomy.

Abbreviations: CI=confidence interval; HR=hazard ratio; IGF-I=insulin-like growth factor; IMD=index of multiple deprivation.

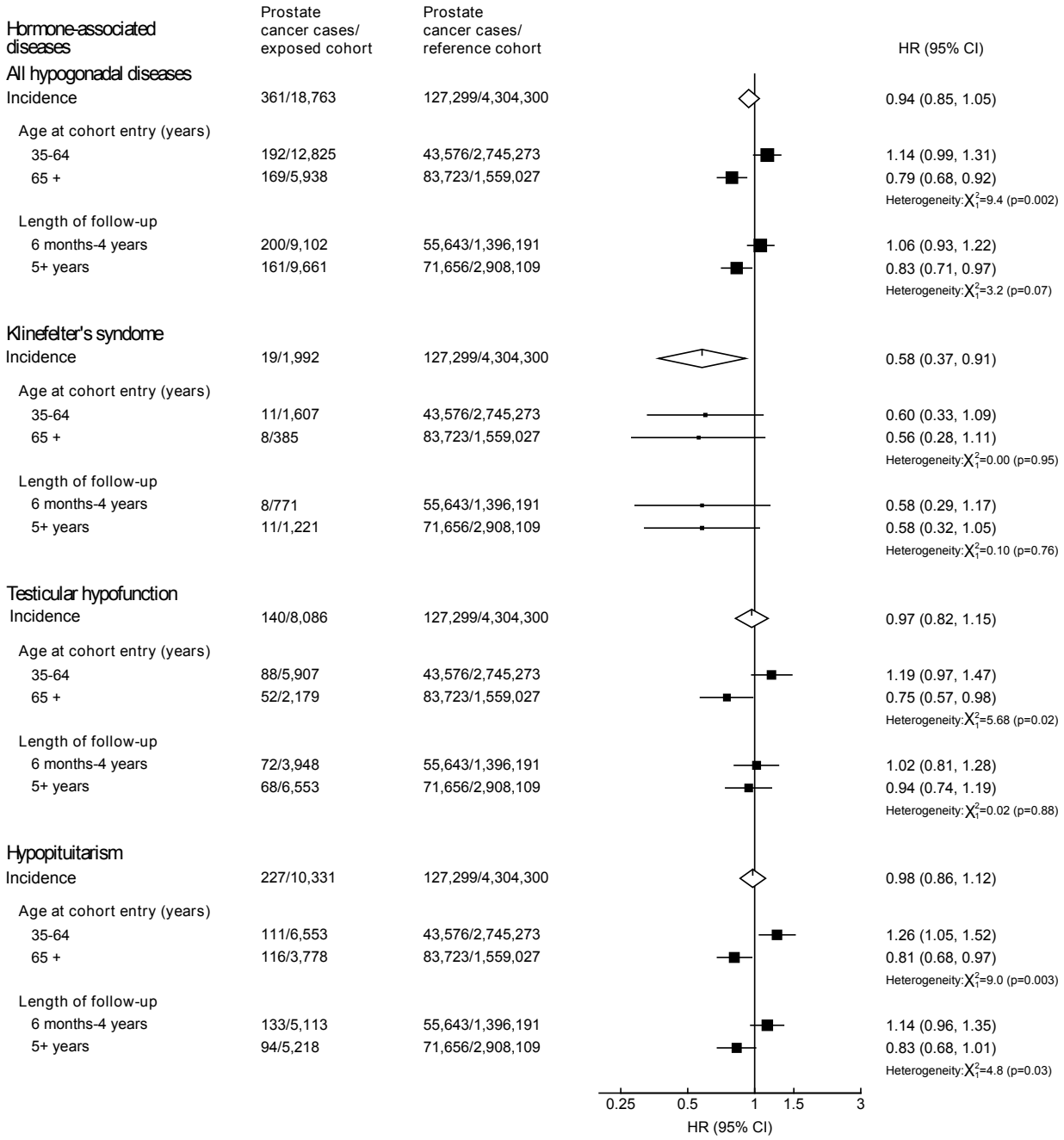


Figure 7.4: Hazard ratio* of prostate cancer diagnosis in men diagnosed with Klinefelter's syndrome, testicular hypofunction and hypopituitarism in comparison with the reference cohort†, stratified by age at cohort entry and time interval.

*HRs adjusted for age (five year groups), year of cohort entry, region of residence (nine regions), IMD rank (quintiles). Results exclude men diagnosed with prostate cancer within the first 6 months of follow-up.

†Conditions used in the reference cohort: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement, and tonsillectomy.

Abbreviations: CI=confidence interval; HR=hazard ratio; IMD=index of multiple deprivation.

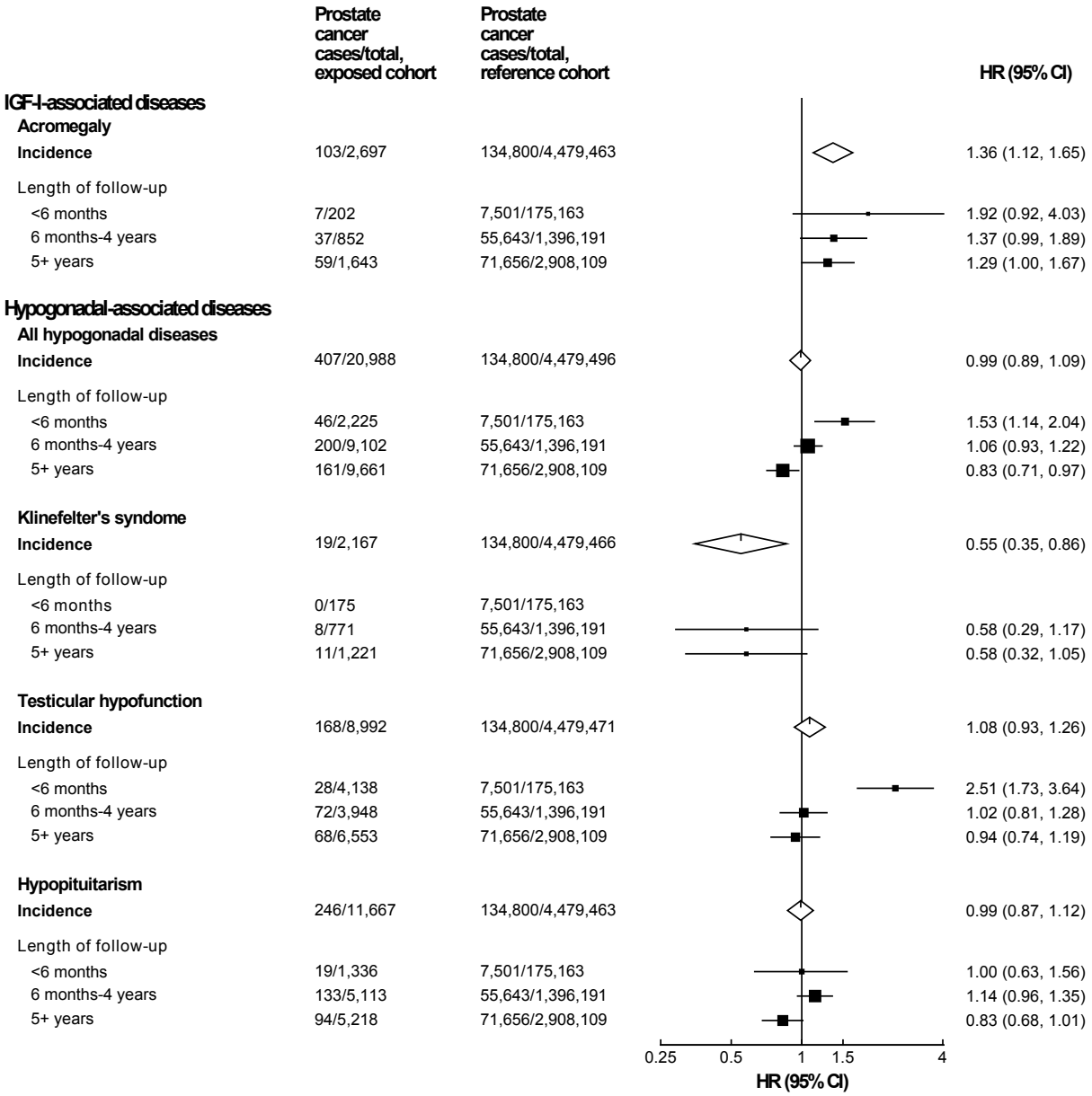


Figure 7.5: Hazard ratio* of prostate cancer incidence by length of follow-up.
*HRs adjusted for age (five year groups), year of cohort entry, region of residence (nine regions), IMD rank (quintiles). Results include men diagnosed with prostate cancer within the first 6 months of follow-up.
†Conditions used in the reference cohort: strabismus, cataract, otitis, varicose veins, haemorrhoids, upper respiratory tract infections, nasal polyps, teeth disorders, inguinal hernia, nail diseases, sebaceous cyst, internal derangement of knee, bunions, vasectomy, dislocations/sprains/strains, bruising, gall bladder disease, appendectomy, hip replacement, knee replacement, and tonsillectomy.
Abbreviations: CI=confidence interval; HR=hazard ratio; IGF-I=insulin-like growth factor; IMD=index of multiple deprivation.

7.4 Discussion

This study of hormone-related diseases and prostate cancer risk is the largest to-date and the first to report a statistically significant association between diagnosis of acromegaly and an increased relative risk of incident prostate cancer, and possibly an increased prostate cancer mortality. Diagnosis with any hypogonadal-associated disease was associated with a reduced risk of prostate cancer mortality, while only Klinefelter's syndrome was associated with a lower risk of prostate cancer diagnosis.

7.4.1 Acromegaly cohort

Acromegaly is associated with clinically high circulating IGF-I concentrations [457, 482–484], which may increase prostate cancer risk by increasing cell division and reducing programmed cell death [215] (Chapter 2). A recent meta-analysis of thirteen studies with data on acromegaly and prostate cancer status reported a 20% increased risk of prostate cancer, however the association was not statistically significant (based on 38 prostate cancer cases) [459]. Individual studies included in the meta-analysis were based on between 25 and 752 men with diagnosed acromegaly, and thus were not able to explore a possible association with prostate cancer mortality. The current analysis, based on 2,495 men diagnosed with acromegaly and 96 prostate cancer cases, found a statistically significant 33% increased risk of prostate cancer incidence and a 44% increased risk of prostate cancer mortality (as the underlying cause of death), although the latter association was not statistically significant, possibly because of the limited number of prostate cancer deaths (n=19).

7.4.2 Hypogonadal cohort

Klinefelter's syndrome, testicular hypofunction and hypopituitarism are conditions which are associated with clinically low circulating testosterone concentrations [465, 468, 474, 475, 485]. Low testosterone concentrations may reduce prostatic androgen receptor signalling, which is important in controlling prostate growth and may affect prostate cancer risk [8–10, 486, 487]. In this analysis, there was no association with prostate cancer diagnosis in the combined hypogonadal cohort, but men diagnosed

with Klinefelter's syndrome had a 42% reduced risk of subsequent prostate cancer diagnosis. This is in agreement with similar findings from previous record linkage studies that have reported a 20-30% relative risk reduction of prostate cancer in men with Klinefelter's syndrome [463, 469, 470]. This is the first study to specifically investigate the associations of testicular hypofunction and hypopituitarism diagnosis with prostate cancer risk and did not find evidence of an association with prostate cancer diagnosis in these men.

The finding that the risk of prostate cancer diagnosis was lower in men with Klinefelter's syndrome may reflect the longer exposure to this genetic condition, or possibly exposure during critical periods [56]. Testicular hypofunction and hypopituitarism generally occur later in life and the underlying condition may be cured, therefore the length of the exposure in men diagnosed with these diseases is likely to be shorter than for men with Klinefelter's syndrome.

Men diagnosed with testicular hypofunction or hypopituitarism had a decreased risk of prostate cancer mortality (in the Klinefelter's syndrome cohort the number of prostate cancer deaths was too low to investigate). Androgen deprivation therapy is a mainstay treatment for advanced prostate cancer [11]. Therefore, the observation that men diagnosed with testicular hypofunction or hypopituitarism had lower prostate cancer mortality, but not overall risk of diagnosis, may indicate that the low androgen environment slows/inhibits prostate tumour growth, reducing the risk of tumour progression and therefore mortality, rather than affecting tumour initiation [211].

7.4.3 Detection bias

Detection bias may play a role in the observed associations. Testosterone replacement therapy (TRT) is a common treatment method for hypogonadal diseases. Although there is no clear evidence that TRT promotes prostate tumorigenesis [488, 489], it is recommended that men receiving TRT are screened and regularly monitored for prostate cancer [481]. Therefore, men receiving TRT may be more likely to attend prostate testing and so be diagnosed with prostate cancer. This screening recommendation may account for the lack of an observed inverse association with prostate cancer diagnosis (observed in Chapter 6), and the heterogeneity in the associations by age observed in the testicular hypofunction and hypopituitarism cohorts, if screening is more common in younger men with these conditions. Although men diagnosed with prostate cancer within the first six months of cohort entry were excluded to reduce detection bias, this bias may not have been eliminated. The observation that men in the Klinefelter's syndrome cohort but not in the testicular hypofunction or hypopituitarism cohorts had a reduced risk of incident prostate cancer may be due to a lower prostate cancer screening attendance in this cohort, possibly related to their higher levels of socio-economic deprivation [44] (which may be related to an increased risk of language disorders and reading disabilities in these men [490, 491]). There are currently no clinical guidelines recommending prostate cancer testing for men with acromegaly [492].

A high proportion of prostate tumours will not become clinically relevant [493]; by examining prostate cancer mortality as the underlying cause of death it is possible to exclude asymptomatic or less aggressive tumours, which may be more likely to be detected due to bias [494]. Therefore, the observed inverse associations with prostate cancer mortality may reflect reduction in clinically significant prostate cancer, which may be less likely to be explained by detection bias. Alternatively, the observed lower prostate cancer mortality might be influenced by the increased PSA testing, and therefore, the increased detection of prostate tumours in these men before and during treatment for hypogonadism. This earlier detection of prostate tumours may lead to a reduction in prostate cancer mortality.

7.4.4 Strengths and limitations

These analyses have some other limitations. Data on prostate tumour characteristics, including stage and grade, were not available. Prostate cancer mortality was used as a proxy for clinical aggressiveness, but the small number of deaths limited the statistical power of analyses particularly for rarer conditions.

Despite choosing diseases that are characterised by aberrant IGF-I and testosterone concentrations, it is not possible to exclude the possibility that other abnormalities or correlates of acromegaly and/or hypogonadal diseases may confound the associations (such as high growth hormone in men diagnosed with acromegaly). However, this may be less likely in the combined hypogonadal disease cohort due to the wide variety of diseases encompassed by Klinefelter's syndrome, testicular hypofunction and hypopituitarism.

Although these diseases are informative proxy measures of aberrant hormone concentrations, no data were available regarding disease treatment which aims to normalise hormone concentrations. Surgery is generally the first-line treatment of acromegaly, but pharmacologic agents are also available which aim to inhibit growth hormone signalling [495]. Long-term biochemical control of the condition is estimated to be achieved in 50-65% of patients who receive either surgical or hormone treatment [458]. TRT is effective in increasing testosterone concentrations in hypogonadal men, but will not treat the underlying cause of the disorder [496]. The proportion of men who are diagnosed with these hormone diseases who will receive treatment is unknown. Hence it is unclear to what extent these men have abnormal hormone concentrations following initial diagnosis. Thus, the strengths of the associations found in this study may be underestimated. This analysis assumes that the men who were diagnosed with these hormone-related diseases and receive treatment will be exposed to the aberrant hormone environment for an adequate duration prior to hormone therapy to elicit an effect.

The investigated diseases were not required to be the principal diagnosis on the hospital record. However, it was only possible to identify men who had hospital inpatient contact; therefore, it is likely that the cohorts did not include all recorded

diagnoses of the exposure diseases or prostate cancer cases. Admission data were not linked to cancer registry records because the permissions were not in place to perform this record linkage. Analysis of the separate UK Biobank dataset showed that approximately 24% of men with cancer registry data did not have a HES APC record of prostate cancer. Therefore, the cohorts may be characterised by men with relatively severe forms of prostate cancer as well as the exposure diseases. Although this may be advantageous in the context of prostate cancer risk, as this may exclude men with less clinically aggressive prostate tumour subtypes, this may bias the results if the extent of missing prostate cancer cases differs between the cohorts. Cohorts identification and "follow-up" used ICD-9 and 10 codes. Coding errors may have led to some misclassification of cohorts and prostate cancer status. This may lead to an underestimation of risk, but any misclassification is likely to be small [497]. This analysis identified prevalent exposure cases based on first recorded hospital diagnosis and is not a follow-up from the first point of onset. Finally, data were not collected on patients who emigrated or received treatment outside England.

Strengths of this analysis include the large size of the study population and the person-based cohort design using linked hospital records from an integrated healthcare system as well as mortality records. It was also possible to adjust for socioeconomic deprivation, and to explore interactions by follow-up time. Although patients are able to opt-out of their data being passed on beyond NHS Digital, the number of opt-outs is low ($\sim$2%) [276]. Therefore, this dataset contains a very high coverage of hospital admissions and death records in England over a nearly 20-year period. This enables the investigation of associations across a variety of rare exposure diseases and subsequent prostate cancer diagnoses and mortality. The very large and heterogeneous reference cohort of 4.3 million men was designed to represent the general population in terms of prostate cancer risk while allowing comparability with the exposed cohorts for the purpose of detailed analyses.

7.5 Conclusion

In summary, using English national hospital admission and mortality records, this study found that men diagnosed with acromegaly had a higher risk of subsequent prostate cancer diagnosis and possibly prostate cancer mortality. This is consistent with the hypothesis that higher IGF-I concentrations may increase prostate cancer risk.

The evidence from this analysis for the possible involvement of testosterone is less clear. Although the diagnosis of a hypogonadal-associated disease was associated with a reduction in prostate cancer mortality, there was only evidence of a relative reduction in prostate cancer diagnosis in the men who were diagnosed with Klinefelter's syndrome. This may indicate that a longer exposure to a low testosterone environment may be required to reduce the risk of prostate cancer, or that detection bias may play a role. The observed lower risk of prostate cancer mortality in men diagnosed with the hypogonadal-associated diseases may support the role of testosterone in prostate cancer progression rather than aetiology.

8

Haematological parameters and prostate cancer risk in UK Biobank

Contents

8.1 Introduction 205
8.1.1 Determinants of red blood cells 205
8.1.2 Determinants of platelets and white blood cells 205
8.1.3 Modifiable risk factors for prostate cancer 205
8.1.4 Aim 206
8.2 Methods 206
8.2.1 UK Biobank 206
8.2.2 Exclusion criteria 210
8.2.3 Statistical analysis 211
8.2.4 Further analyses 213
8.2.5 Statistical software 214
8.3 Results 215
8.3.1 Correlation between blood indices 219
8.3.2 Associations between haematological indices and prostate cancer incidence and mortality 221
8.3.3 Further analyses 229
8.4 Discussion 236
8.4.1 Associations with red blood cell indices 236
8.4.2 Associations with platelet and white blood cell indices . 236
8.4.3 Diet and other factors 237
8.4.4 Reverse causation 237
8.4.5 Strengths and limitations 238
8.4.6 Conclusion 239

8.1 Introduction

Risk factors for prostate cancer are not well understood (Chapter 2). Haematological indices are associated with a range of exposures, therefore may be useful surrogate markers to identify possible disease risk factors.

8.1.1 Determinants of red blood cells

Testosterone, folate, vitamin B_{12} and iron are fundamental to the generation of red blood cells and low concentrations of any of these factors can lead to changes in red blood cell counts and morphologies [498–505].

Testosterone is thought to stimulate erythropoietin production and acts directly on the bone marrow to increase the number of erythropoietin-responsive cells [506]. It is widely observed that hypogonadal men have lower haemoglobin and hematocrit than eugonadal men, and treatment for hypogonadism will increase haematocrit [499, 500].

Folate, vitamin B_{12} and iron are nutrients that are also crucial for the generation of red blood cells. Erythroblasts (young red blood cells) are dependent on folate and vitamin B_{12} for proliferation, deficiencies in these vitamins will impair DNA synthesis and lead to erythroblast apoptosis [501]. Erythroblasts are also dependent on iron to synthesise haemoglobin, which is responsible for the transportation of oxygen around the body [501].

8.1.2 Determinants of platelets and white blood cells

Inflammation is defined as a response to stimulation from pathogens or endogenous signals such as damaged cells. Chronic inflammation can also be caused by obesity, arthritis and cardiovascular disease [507]. Inflammation and infection directly stimulates increases in platelet and white blood cell production [505, 508].

8.1.3 Modifiable risk factors for prostate cancer

Analyses in Chapters 6 and 7 indicated that low testosterone concentrations may reduce prostate cancer risk. Additionally, the nutritional factors folate, vitamin B_{12}

and iron may also be associated with prostate cancer risk, but these associations are inconsistent [137, 139, 509–511] (Chapter 2). Similarly, previous studies have indicated that chronic inflammation and/or the immune response may be associated with cancer overall incidence [512, 513], but it is unclear whether these factors are associated with prostate cancer risk [119–121] (Chapter 2).

Although some studies have examined the associations between a limited number of blood parameters and overall cancer risk [514–518], few have had the power to examine these exposures in relation to prostate cancer diagnosis or mortality [514, 515]. Moreover, these studies have relied on a single measure at study baseline, which may lead to an attenuation of risk estimates [290].

UK Biobank is a large prospective cohort including approximately 220,000 men, currently with ∼6,000 incident cases of prostate cancer and 350 prostate cancer deaths. At baseline, complete blood count and reticulocyte indices were measured across the entire cohort, and repeat measurements were made in ∼9,000 men. The large size of this dataset, repeat measurements of biomarkers and breadth of the haematological parameters measured provides a unique opportunity to investigate possible risk factors for prostate cancer.

8.1.4 Aim

In order to enhance the understanding of risk factors for prostate cancer, this chapter aims to use haematological indices as possible surrogate markers of exposures including testosterone, dietary factors and inflammation/infection, to examine their associations with prostate cancer risk.

8.2 Methods

8.2.1 UK Biobank

Study Design

UK Biobank is a large prospective cohort study designed to be a resource for research into the causes of disease. Further details of the study protocol and data

collection can be found in Chapter 3. This research has been conducted using the UK Biobank resource under application number 3282.

In brief, all participants were registered with the 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 via postal invitation. Overall, 503,317 men and women aged 40-69 years consented to join the cohort and attended one of 22 assessment centres throughout England, Wales and Scotland between 2006-2010, resulting in a participation rate of 5.5% [519].

Baseline assessment and dietary follow-up

At baseline participants provided data on a range of sociodemographic, physical, lifestyle, and health-related factors and summary diet information (including supplement and medication use) via self-completed touch-screen questionnaires and a computer assisted personal interview [519]. Body mass index (BMI) and other covariates were measured at the assessment centre.

Blood samples were taken from all participants except for a small proportion who declined, were deemed unable to, or where the attempt was abandoned for either technical or health reasons (0.3%) [519].

To prevent the blood from clotting it was collected into Ethylenediaminetetraacetic acid (EDTA) vacutainers and shipped to the central processing laboratory in temperature-controlled shipping boxes (at 4 °C) [280]. On arrival in the central laboratory, the blood components were measured using a Beckman automated haematology analyser, and reticulocyte parameters were measured using a COULTER LH 750 System, typically within 24 hours of blood draw [280]. A maximum of 31 haematological parameters were generated, including a mixture of measured, derived and calculated values [281]. Descriptions of the blood parameters included in this analysis and the formulae to estimate these measures are displayed in Table 8.1.

For red blood cell and platelet parameters the values analysed in this chapter are as measured by UK Biobank. For subtypes of white blood cells (neutrophils,

eosinophils, basophils, monocytes, lymphocytes), differential parameters (the proportion of specific white cells in the white blood cell population) measured by the machine were used to recalculate counts of each cell type.

Table 8.1: Summary description of haematological parameters and measurement/calculation methods by UK Biobank.

Haematological parameter	Description	Measurement/calculation
Red blood cells		
RBC	Number of red blood cells in the sample	Measured
RCDW	Spread of red blood cell population	Derived*
HCT	Volume occupied by red blood cells in the blood	(MCV x RBC)/10
HGB	Total haemoglobin concentration in sample	Measured
MCV	Average volume of red blood cells	Derived*
MCH	Mass of haemoglobin in the average red blood cell	(HGB / RBC) x 10
MCHC	Average mass of haemoglobin per the relative volume of red blood cells in the whole blood sample	(HGB / HCT) x 100
MSCV	Average volume of all red blood cells (both mature and reticulocytes)	Derived*
RET	Number of reticulocytes in the red blood cell sample	% Reticulocyte x RBC
MRV	Average volume of individual reticulocytes in the sample	Calculation unknown
IMF	Ratio of the count of the highest light scatter reticulocytes to the total count of reticulocytes	High light scatter reticulocytes / RET
Platelets		
PLT	Number of platelets in the sample	Measured
PCT	Volume occupied by platelets in the blood	(PLT × MPV) / 10,000
PDW	Variation in platelet volume	Derived*
MPV	Average volume of individual platelets in the sample	Derived*
White blood cells		
WBC	Number of white blood cells in the sample	Measured
Neutrophil count	Number of neutrophils in the white blood cell sample	(% Proportion of neutrophils/100) x WBC
Eosinophil count	Number of eosinophils in the white blood cell sample	(% Proportion of eosinophils/100) x WBC
Basophil count	Number of basophils in the white blood cell sample	(% Proportion of basophils/100) x WBC
Monocyte count	Number of monocytes in the white blood cell sample	(% Proportion of monocytes/100) x WBC
Lymphocyte count	Number of lymphocytes in the white blood cell sample	(% Proportion of lymphocytes/100) x WBC

Measurement/calculation methods were obtained from: `http://biobank.ctsu.ox.ac.uk/crystal/docs/haematology.pdf`.
*Derived values were calculated within the instrument using multiple scatterplots and histograms for each sample.
Abbreviations: HCT=haematocrit; HGB=haemoglobin concentration; IMF=immature reticulocyte fraction; MCH=mean corpuscular haemoglobin; MCHC=mean corpuscular haemoglobin concentration; MCV=mean corpuscular volume; MPV=mean platelet volume; MRV=mean reticulocyte volume; MSCV=mean sphered corpuscular volume; PCT=plateletcrit; PDW=platelet distribution width; PLT=platelet count; RBC=red blood cell count; RCDW=red blood cell distribution width; RET=reticulocyte count; WBC=white blood cell count.

Repeat assessment and follow-up

Repeat assessment data and biological samples were collected in a subset of participants (∼9,000 men) between August 2012 and June 2013 at the UK Biobank Co-ordinating Centre. Participants who lived within a 35 km radius were invited to attend with an overall response rate of 21% [281, 291].

Cancer registration data were provided via record linkage to the NHS Central Register, until the censoring date (31st March 2016 in England and Wales and 31st 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 31st January 2018 and 30th November 2016, respectively).

In the analysis of incident prostate cancer, the endpoint was defined as the first diagnosis of prostate cancer, or death from prostate cancer (International Classification of Diseases Tenth revision code (ICD-10) C61) [478], whichever was recorded first. The prostate cancer cases which were identified via death records were included only if death from prostate cancer preceded the NHS Central Register censoring date. In the analysis of prostate cancer mortality, the endpoint was defined 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)) [478], death, or censoring date, whichever occurred first.

8.2.2 Exclusion criteria

There were 209,686 men included in this analysis (Figure 8.1), after excluding 9,871 men with prevalent cancer (except C44: non-melanoma skin cancer), 56 withdrew consent, 213 participants were excluded as they were identified as being genetically female [520], and 9,364 men did not have blood data available.

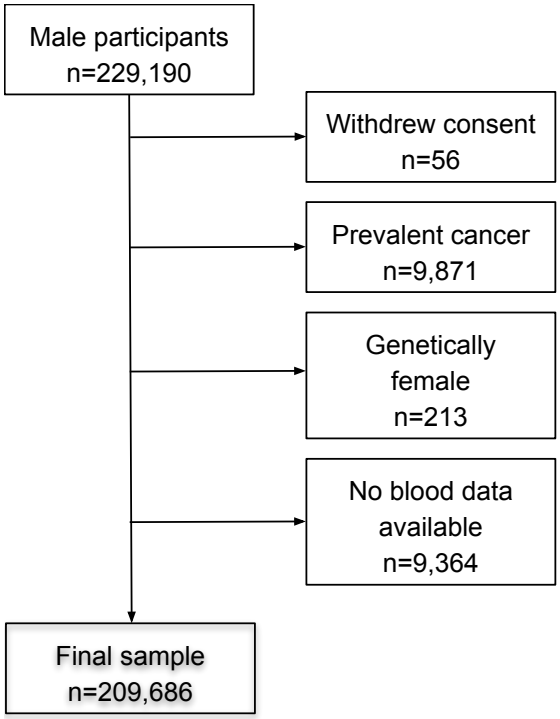


Figure 8.1: Participant selection chart

8.2.3 Statistical analysis

Partial correlations between the blood indices were estimated after adjustment for region (10 UK regions), age at recruitment (<45, 45–49, 50–54, 55–59, 60–64, and ≥65 years), BMI (<25, ≥25-<30, ≥30-<35, ≥35 kg/m^2, unknown (0.4%)).

Hazard ratios (HRs) and 95% confidence intervals (CIs) of prostate cancer diagnosis and mortality were estimated using Cox proportional hazards models, using attained age as the underlying time variable. Analyses were stratified by geographic region of recruitment, age at recruitment (categories as defined above), and adjusted for Townsend deprivation score (fifths, unknown (0.1%)), racial/ethnic group (white, mixed background, Asian, black, other, and unknown (0.6%)), height (<170, ≥170–<175, ≥175–<180, ≥180 cm, and unknown (0.4%)), lives with a wife or partner (no, yes), BMI (categories as defined above), 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-

≤9, ≥10-<20, ≥20 g ethanol/day, unknown (0.5%)), and diabetes (no, yes, and unknown (0.5%)). Categories of the adjustment covariates were defined *a priori* based on previous analyses [74].

The blood parameters investigated were chosen *a priori* on the basis of data availability: i) red blood cells: red blood cell (erythrocyte) count, red blood cell distribution width, haematocrit, haemoglobin concentration, mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), mean sphered cell volume (MSCV), reticulocyte count, reticulocyte volume, immature reticulocyte fraction; ii) platelets: platelet count, plateletcrit, platelet distribution width, mean platelet (thrombocyte) volume; iii) white blood cells: white blood cell (leukocyte) count, neutrophil count, eosinophil count, basophil count, monocyte count and lymphocyte count.

Due to measurement error and within person variability, using single measures at baseline may lead to substantial under-estimation of potential associations with disease risk (i.e. regression dilution bias) [290]. UK Biobank repeated the blood measurements during follow-up in a sub-sample of 8,667 men. HRs for trend were estimated per one unit increase in standard deviation (SD) and were corrected for regression dilution bias using the McMahon-Peto ratio [521]. In the categorical analysis, haematological data were categorised into fifths of the distribution of the whole cohort at study baseline, with the 1st fifth representing the lowest parameter values and the 5th representing the highest. 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 in all groups except the reference group) and used to obtain group-specific 95% CIs, which enabled comparisons across different exposure categories [430].

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

8.2.4 Further analyses

Tests for heterogeneity

To examine whether associations with prostate cancer risk differed for cancers diagnosed shortly after recruitment or in men who were diagnosed at a younger age, heterogeneity in the associations of haematological indices with incident prostate cancer was tested by time from recruitment to diagnosis (<5 and ≥ 5 years) and age at diagnosis (<65 and ≥ 65 years). Stratified Cox models were fitted based on competing risks and heterogeneity in the risk coefficients and standard errors in the two subgroups were tested using the χ^2 test of heterogeneity. Heterogeneity by subgroups in the associations with prostate cancer mortality was not tested due to the relatively small numbers of deaths and therefore limited statistical power.

For non-case defined chacteristics, heterogeneity in the associations with incident prostate cancer by age at blood collection (<60; ≥ 60 years), BMI (<30; ≥ 30 kg/m^2), smoking (never; ever smokers) and diabetes status (yes; no) was tested using the χ^2 test of interaction between subgroups.

Subgroup cut-points were defined *a priori* based on the data availability.

Sensitivity analysis

Family history of prostate cancer was not included as an adjustment factor in the primary analysis due to the limited availability of this data (45.1% missing/unknown). It is likely that this information is not missing at random; men with a family history of prostate cancer are more likely to be able to recall this information. Family history of prostate cancer was further adjusted for as a sensitivity analysis (no, yes (brother or father), and unknown). To adjust for the possibility of laboratory drift, analyses were repeated following additional adjustment for year of blood collection (continuous).

The primary analysis was repeated after: i) excluding men who reported that they were regularly taking folic acid, folate or multivitamin supplementation at baseline (n=41,323); ii) excluding men who reported that they were regularly taking prescription hydroxyurea, interferon, heparin, erythropoietin, anaemia treatment,

immunosuppressants, antihistamines or anti-inflammatories at baseline (n=64,284), as these medications are designed to alter blood parameter values; and iii) after removing outliers (outside the range of [lower quartile – 3*interquartile range, upper quartile + 3*interquartile range]) (n=18-1,940 for immature reticulocyte fraction and red cell distribution width, respectively).

8.2.5 Statistical software

All statistical tests were two-sided and owing, to the large number of tests conducted, with the statistical significance threshold set at $p<0.01$. Data analyses used Stata Statistical Software release 14.1 (Stata Corp., College Station, TX, USA), and figures were created in R version 3.2.3.

8.3 Results

209,686 men were included in this analysis and, after a mean follow-up of 6.8 years (SD=1.3 years) 5,723 (2.7%) were diagnosed with prostate cancer and 323 died from the disease. Table 8.2 summarises the baseline characteristics of study participants. Mean age at recruitment was 56.6 years (SD=8.2), and men had a mean BMI of 27.8 kg/m^2. 12% of men were current smokers, 43% reported drinking $\geq$20 g alcohol per day. 28% reported having a PSA test prior to baseline and 7.5% had a family history of prostate cancer.

Median and interquartile range values for haematological data are displayed in Table 8.2. Characteristics of men by red blood cell, platelet, and white blood cell counts by 1st, 3rd and 5th fifths at baseline are displayed in Table 8.3. Men with higher red blood cell counts were younger, had a higher BMI and were less likely to smoke. Men with higher platelet counts were younger and more likely to smoke. Men with higher white blood cell counts were older, had a higher BMI and were more likely to smoke (Table 8.3). Regression dilution ratios are displayed in Table 8.4. Basophil count and MCHC had the lowest regression dilution ratios (0.20 and 0.24, respectively). Regression dilution ratios for the other haematological parameters ranged between 0.44 - 0.83 (Table 8.4).

Table 8.2: Baseline characteristics and blood indices of men who remained free from prostate cancer and men who developed prostate cancer in UK Biobank.

	Men free from prostate cancer (n=203,963)	Men who developed prostate cancer (n=5,723)
Sociodemographic		
Age at recruitment (years), mean (SD)	56.4 (8.2)	62.2 (5.2)
Most deprived fifth, % (n)	19.9 (40,618)	15.7 (896)
Black ethnicity, % (n)	1.5 (2,966)	2.0 (117)
Not in paid/self-employment, % (n)	38.2 (77,911)	57.1 (3,268)
Living with partner, % (n)	76.3 (155,531)	79.2 (4,533)
Anthropometric, mean (SD)		
Height (cm)	175.7 (6.8)	175.1 (6.7)
BMI (kg/m^2)	27.8 (4.2)	27.6 (3.8)
Lifestyle, % (n)		
Current cigarette smokers	12.5 (25,537)	9.2 (524)
Drinking alcohol ≥20 g per day	43.3 (88,402)	42.5 (2,432)
Low physical activity (0-10 METs per week)	27.5 (56,017)	26.2 (1,500)
Health history, % (n)		
Hypertension	52.0 (106,004)	59.0 (3,376)
Diabetes	6.9 (14,092)	6.0 (342)
Prostate specific factors, % (n)		
Ever had a PSA test	27.3 (55,655)	45.9 (2,624)
Family history of prostate cancer	7.4 (15,005)	13.0 (745)
Blood measures at initial assessment, median (interquartile range)		
Red blood cell		
Red blood cell count (10^{12} cells/L)	4.75 (4.52 - 4.99)	4.73 (4.49 - 4.98)
Red blood cell distribution width (%)	13.3 (12.9 - 13.8)	13.4 (13.0 - 13.9)
Haematocrit (%)	43.4 (41.5 - 45.2)	43.3 (41.4 - 45.3)
Haemoglobin concentration (g/dL)	15.0 (14.4 - 15.7)	15.0 (14.3 - 15.6)
Mean corpuscular volume (pg)	91.4 (88.8 - 94.1)	91.5 (88.9 - 94.2)
Mean corpuscular haemoglobin (pg)	31.7 (30.7 - 32.7)	31.7 (30.7 - 32.6)
Mean corpuscular haemoglobin concentration (g/dL)	34.6 (34.0 - 35.2)	34.5 (33.9 - 35.2)
Mean sphered cell volume (fL)	82.5 (79.2 - 85.9)	82.6 (79.4 - 86.1)
Reticulocyte count (10^{12} cells/L)	0.06 (0.05 - 0.08)	0.06 (0.05 - 0.08)
Mean reticulocyte volume (fL)	106 (102 - 111)	106 (102 - 111)
Immature reticulocyte fraction	0.29 (0.25 - 0.33)	0.29 (0.25 - 0.33)
Platelet		
Platelet count (10^9 cells/L)	234 (202 - 269)	233 (200 - 269)
Plateletcrit (%)	0.22 (0.19 - 0.24)	0.21 (0.19 - 0.24)
Platelet distribution width (%)	16.5 (16.2 - 16.9)	16.5 (16.2 - 16.9)
Mean platelet volume (fL)	9.18 (8.54 - 9.90)	9.17 (8.53 - 9.93)
White blood cell		
White blood cell count (10^9 cells/L)	6.69 (5.67 - 7.90)	6.70 (5.70 - 7.89)
Neutrophil count (10^9 cells/L)	4.06 (3.29 - 5.00)	4.15 (3.36 - 5.09)
Eosinophil count (10^9 cells/L)	0.15 (0.10 - 0.24)	0.15 (0.09 - 0.24)
Basophil count (10^9 cells/L)	0.029 (0.019 - 0.044)	0.028 (0.019 - 0.043)
Monocyte count (10^9 cells/L)	0.49 (0.39 - 0.61)	0.50 (0.40 - 0.62)
Lymphocyte count (10^9 cells/L)	1.81 (1.47 - 2.21)	1.78 (1.45 - 2.18)

Abbreviations: BMI=body mass index; MET=metabolic equivalent; PSA=prostate-specific antigen, SD=standard deviation.

Table 8.3: Baseline characteristics of 209,686 men in UK Biobank according to red blood cell, platelet and white blood cell counts at initial assessment.

	Fifths of red blood cell count			Fifths of platelet cell count			Fifths of white blood cell count		
	1	3	5	1	3	5	1	3	5
No. of men	41,446	41,786	41,034	41,102	41,230	41,157	41,325	41,892	41,266
Age at recruitment*, yrs	58.5 (7.7)	56.3 (8.2)	55.1 (8.4)	57.9 (8.1)	56.5 (8.2)	55.6 (8.1)	55.2 (8.2)	56.8 (8.1)	57.4 (8.1)
BMI*, kg/m^2	27.1 (4.4)	27.8 (4.1)	28.6 (4.2)	28.0 (4.3)	27.8 (4.2)	27.8 (4.3)	26.7 (3.7)	27.9 (4.0)	28.8 (4.9)
Height*, cm	175.2 (6.9)	175.9 (6.8)	175.5 (6.9)	176.2 (6.9)	175.8 (6.8)	174.7 (6.8)	176.7 (6.8)	175.6 (6.8)	174.5 (6.8)
Smoking, n (%)									
Never	17,891 (42.6)	21,231 (49.9)	22,203 (53.8)	20,969 (50.0)	20,816 (49.7)	18,965 (45.6)	24,491 (58.2)	21,593 (50.9)	14,856 (35.6)
Former	17,686 (42.1)	16,128 (37.9)	14,012 (33.9)	16,087 (38.4)	15,962 (38.1)	15,834 (38.1)	15,099 (35.9)	16,675 (39.3)	15,175 (36.4)
Current	6,199 (14.7)	4,981 (11.7)	4,779 (11.6)	4,618 (11.0)	4,874 (11.6)	6,531 (15.7)	2,283 (5.4)	3,909 (9.2)	11,320 (27.1)
Educational level, n (%)									
No degree	7,466 (17.8)	8,034 (18.9)	7,993 (19.4)	7,545 (18.0)	7,831 (18.7)	8,059 (19.4)	7,520 (17.9)	7,915 (18.7)	8,220 (19.7)
Degree	25,817 (61.4)	29,400 (64.1)	27,963 (62.6)	26,592 (63.4)	26,816 (63.9)	25,283 (60.3)	28,971 (68.9)	27,044 (63.8)	23,223 (55.7)
Physical activity, n (%)									
Low	10,684 (25.4)	11,448 (26.9)	12,680 (30.7)	11,228 (26.8)	11,308 (27.0)	12,124 (29.1)	9,644 (22.9)	11,480 (27.1)	13,588 (32.6)
Moderate	20,036 (47.7)	20,606 (48.4)	19,052 (46.1)	20,485 (48.8)	20,147 (48.1)	19,116 (45.9)	21,429 (50.9)	20,422 (48.1)	18,190 (43.6)
High	9,991 (23.8)	9,247 (21.7)	7,958 (19.3)	8,958 (21.4)	9,186 (21.9)	8,874 (21.3)	9,886 (23.5)	9,215 (21.7)	8,295 (19.9)
Ethnicity, n (%)									
White	40,320 (95.9)	40,481 (96.3)	36,815 (87.6)	38,949 (92.9)	39,514 (94.2)	39,326 (93.8)	38,872 (92.4)	40,164 (94.7)	39,330 (94.3)
Not white	1,485 (3.5)	1,870 (4.4)	4,207 (10.2)	2,714 (6.5)	2,164 (5.2)	2,057 (5.0)	2,964 (7.0)	2,026 (4.8)	2,137 (5.1)
Diabetes at baseline, n (%)									
No	37,762 (89.8)	39,862 (93.6)	38,456 (93.1)	38,097 (90.8)	39,100 (93.2)	38,559 (91.9)	40,243 (95.7)	39,510 (93.1)	36,720 (88.0)
Yes	4,089 (9.7)	2,526 (5.9)	2,544 (6.2)	3,601 (8.6)	2,625 (6.3)	2,812 (6.8)	1,652 (3.9)	2,700 (6.4)	4,727 (11.3)
Marital status, n (%)									
Married	31,181 (73.9)	32,935 (77.0)	31,388 (75.5)	32,116 (76.6)	32,328 (77.1)	30,742 (73.3)	32,689 (77.4)	32,865 (77.1)	29,637 (70.5)
Not married	10,853 (25.7)	9,638 (22.5)	9,906 (23.8)	9,822 (23.4)	9,565 (22.8)	10,866 (25.9)	9,370 (22.2)	9,553 (22.4)	12,092 (28.8)

Percentages may not match due to missing data. The 1st fifth represents the lowest parameter value and the 5th fifth the highest.
*Values are means (SD).
Abbreviations: BMI=body mass index; SD=standard deviation.

Table 8.4: Median blood indices at initial assessment and first resurvey and regression dilution ratios.

Blood parameters	Baseline (n=8,698)	Resurvey (n=8,698)	Regression dilution ratio*
Red blood cell, median (interquartile range)			
Red blood cell count (10^{12} cells/L)	4.72 (4.49 - 4.96)	4.74 (4.50 - 4.97)	0.78
Red blood cell distribution width (%)	13.3 (12.9 - 13.8)	13.3 (12.9 - 13.8)	0.68
Haematocrit (%)	43.2 (41.3 - 45.1)	43.1 (41.2 - 44.9)	0.66
Haemoglobin concentration (g/dL)	15.0 (14.3 - 15.6)	15.1 (14.4 - 15.7)	0.71
Mean corpuscular volume (fL)	91.6 (89.0 - 94.2)	90.9 (88.5 - 93.7)	0.79
Mean corpuscular haemoglobin (pg)	31.7 (30.8 - 32.7)	31.8 (30.8 - 32.8)	0.82
Mean corpuscular haemoglobin concentration (g/dL)	34.6 (34.0 - 35.3)	35.3 (34.0 - 35.8)	0.24
Mean sphered cell volume (fL)	82.7 (79.5 - 86.1)	86.9 (83.9 - 90.3)	0.70
Reticulocyte count (10^{12} cells/L)	0.06 (0.04 - 0.08)	0.05 (0.04 - 0.07)	0.53
Mean reticulocyte volume (fL)	106 (102 - 111)	112 (107 - 116)	0.58
Immature reticulocyte fraction	0.29 (0.25 - 0.33)	0.31 (0.27 - 0.35)	0.44
Platelet, median (interquartile range)			
Platelet count (10^9 cells/L)	232 (201 - 267)	216 (185 - 250)	0.77
Plateletcrit (%)	0.21 (0.19 - 0.24)	0.19 (0.17 - 0.22)	0.71
Platelet distribution width (%)	16.5 (16.2 - 16.9)	16.6 (16.3 - 17.0)	0.66
Mean platelet (thrombocyte) volume (fL)	9.20 (8.53 - 9.90)	9.00 (8.34 - 9.70)	0.83
White blood cell, median (interquartile range)			
White blood cell count (10^9 cells/L)	6.50 (5.50 - 7.60)	6.64 (5.70 - 7.80)	0.67
Basophil count (10^9 cells/L)	0.03 (0.02 - 0.04)	0.03 (0.02 - 0.04)	0.20
Eosinophil count (10^9 cells/L)	0.15 (0.09 - 0.23)	0.15 (0.09 - 0.23)	0.71
Lymphocyte count (10^9 cells/L)	1.77 (1.44 - 2.15)	1.75 (1.41 - 2.16)	0.80
Monocyte count (10^9 cells/L)	0.48 (0.39 - 0.60)	0.46 (0.35 - 0.60)	0.60
Neutrophil count (10^9 cells/L)	3.92 (3.20 - 4.81)	4.12 (3.36 - 5.04)	0.61

*Regression dilution ratios were estimated using the MacMahon-Peto method [521].

8.3.1 Correlation between blood indices

After adjusting for region, age and BMI, measures of mature red blood cell size were generally positively correlated (Figure 8.2). Haematocrit had a strong positive partial correlation with haemoglobin concentration (r=0.91). MCH was positively correlated with MCV, MCHC, MSCV (r=0.84, 0.56 and 0.55, respectively). Haematocrit was positively correlated with haemoglobin concentration and red blood cell count (r=0.91 and 0.79, respectively).

Measures of platelet volume were positively correlated (Figure 8.2). Platelet count was positively correlated with neutrophil and monocyte counts (r=0.24 and 0.13, respectively).

White blood cell counts were positively correlated with each other (Figure 8.2). White blood cell count was positively correlated with neutrophil and lymphocyte counts (r=0.85 and 0.61, respectively).

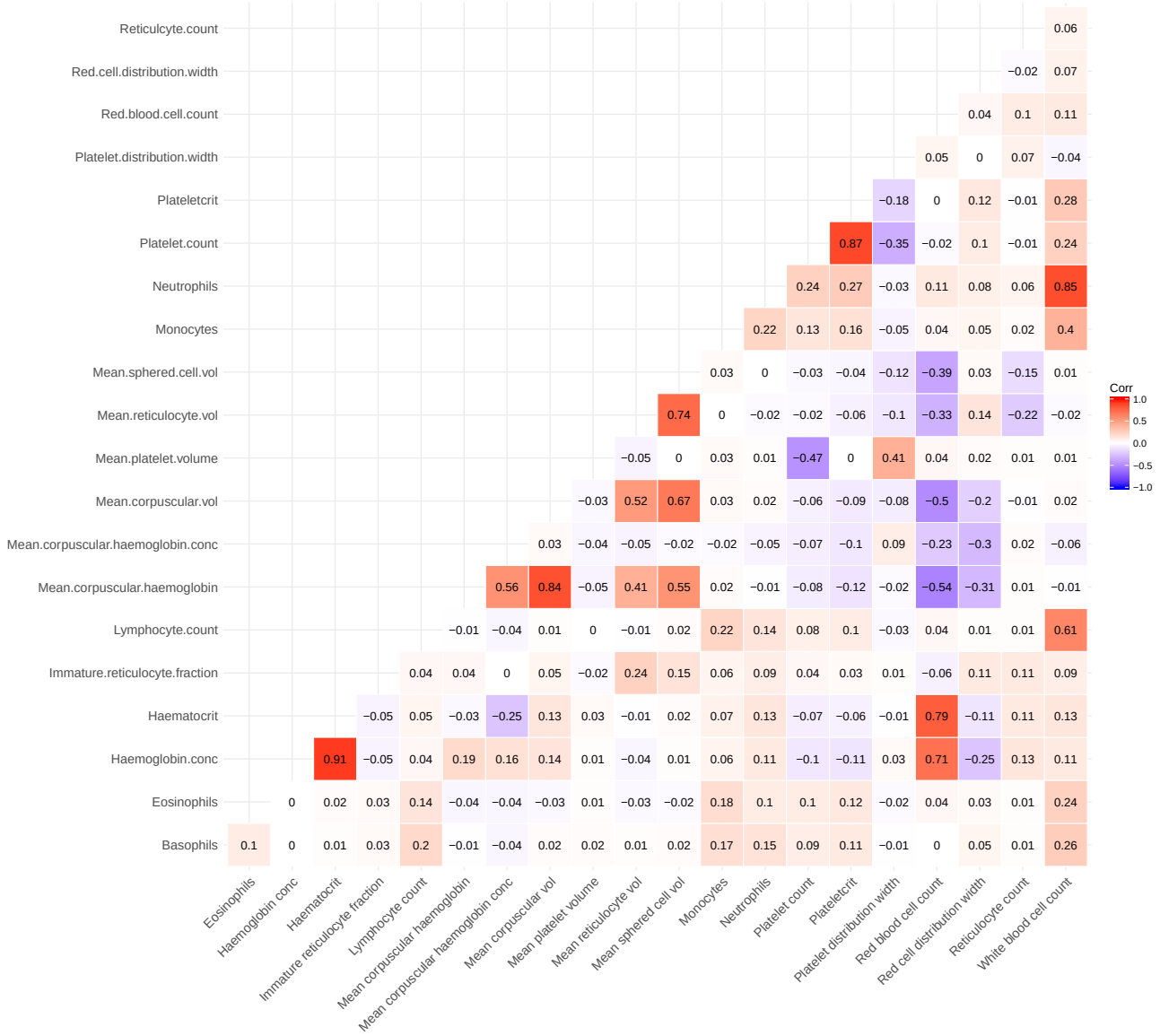


Figure 8.2: Partial correlation matrix of haematological indices.
Correlations are adjusted for region (10 regions), age (<45, 45–49, 50–54, 55–59, 60–64, and ≥65 years) and BMI (<25, ≥25-<30, ≥30-<35, ≥35 kg/m^2, unknown).
Abbreviations: BMI=body mass index.

8.3.2 Associations between haematological indices and prostate cancer incidence and mortality

Higher total red blood cell count was associated with a 9% increased prostate cancer risk per 1 SD increase (95% CI 1.05-1.13; $P_{trend}<0.001$). Higher MCV (HR per 1 SD increase=0.90, 95% CI 0.87-0.93; $P_{trend}<0.001$), MCH (HR per 1 SD increase=0.90, 95% CI 0.87-0.93; $P_{trend}<0.001$), MCHC (HR per 1 SD increase=0.87, 95% CI 0.77-0.97; $P_{trend}=0.01$), MSCV (HR per 1 SD increase=0.91, 95% CI 0.87-0.94; $P_{trend}<0.001$) and mean reticulocyte volume (HR per 1 SD increase=0.87, 95% CI 0.83-0.91; $P_{trend}<0.001$) were associated with a lower prostate cancer risk (Figure 8.3). Red blood cell indices were not associated with prostate cancer mortality (Figure 8.4).

Higher platelet count and plateletcrit were associated with an increased risk of prostate cancer diagnosis (HR per 1 SD increase=1.07, 95% CI 1.04-1.11; $P_{trend}<0.001$; and HR per 1 SD increase=1.10, 95% CI 1.06-1.14, $P_{trend}<0.001$, respectively) (Figure 8.5). Platelet indices were not associated with prostate cancer mortality (Figure 8.6).

White blood cell counts were not associated with prostate cancer diagnosis (Figure 8.7), but higher white blood cell count and neutrophil count were associated with an increased risk of prostate cancer mortality (HR per 1 SD increase=1.14, 95% CI 1.05-1.24; $P_{trend}=0.002$; and HR per 1 SD increase=1.27, 95% CI 1.09-1.48, $P_{trend}=0.002$, respectively) (Figure 8.8).

Estimates are corrected for regression dilution bias; uncorrected results are displayed in Table 8.5.

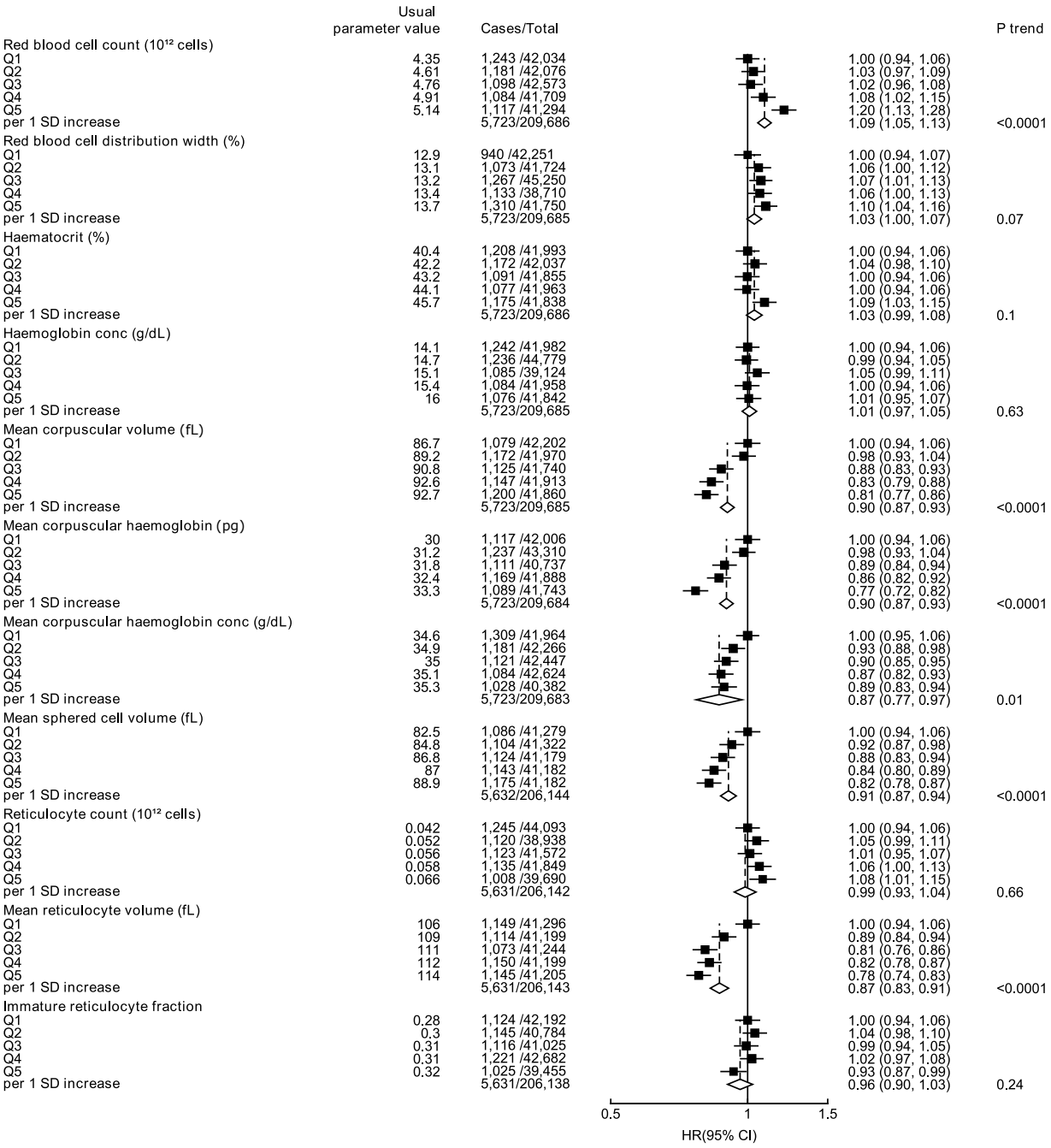


Figure 8.3: Hazard ratio (95% CIs) of prostate cancer diagnosis by fifths of red blood cell indices.

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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). HRs per 1 SD increase are adjusted for regression dilution bias.

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio; IQR=interquartile range; SD=standard deviation.

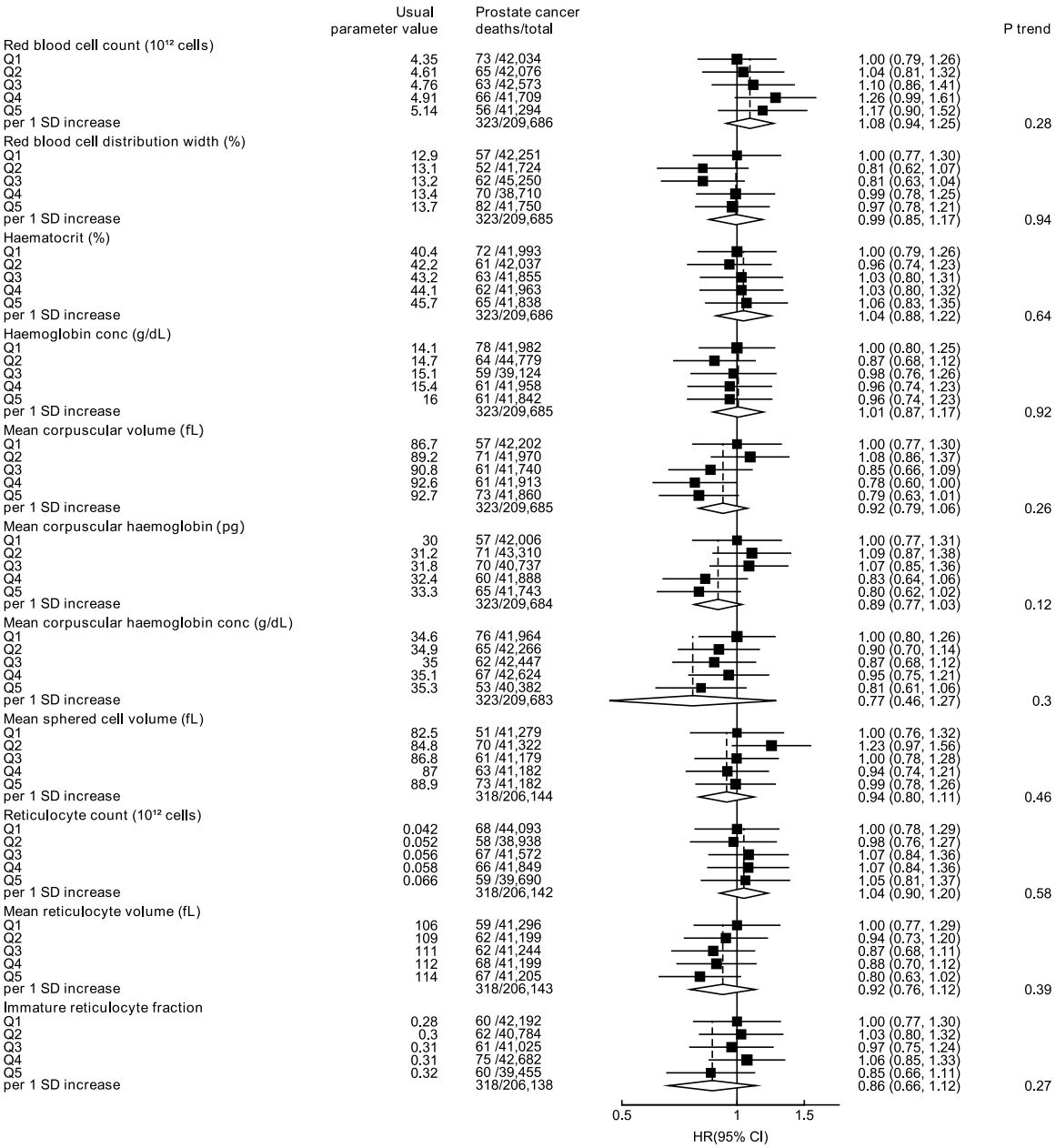


Figure 8.4: Hazard ratio (95% CIs) of prostate cancer mortality by fifths of red blood cell indices.

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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). HRs per 1 SD increase are adjusted for regression dilution bias.

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio; IQR=interquartile range; SD=standard deviation.

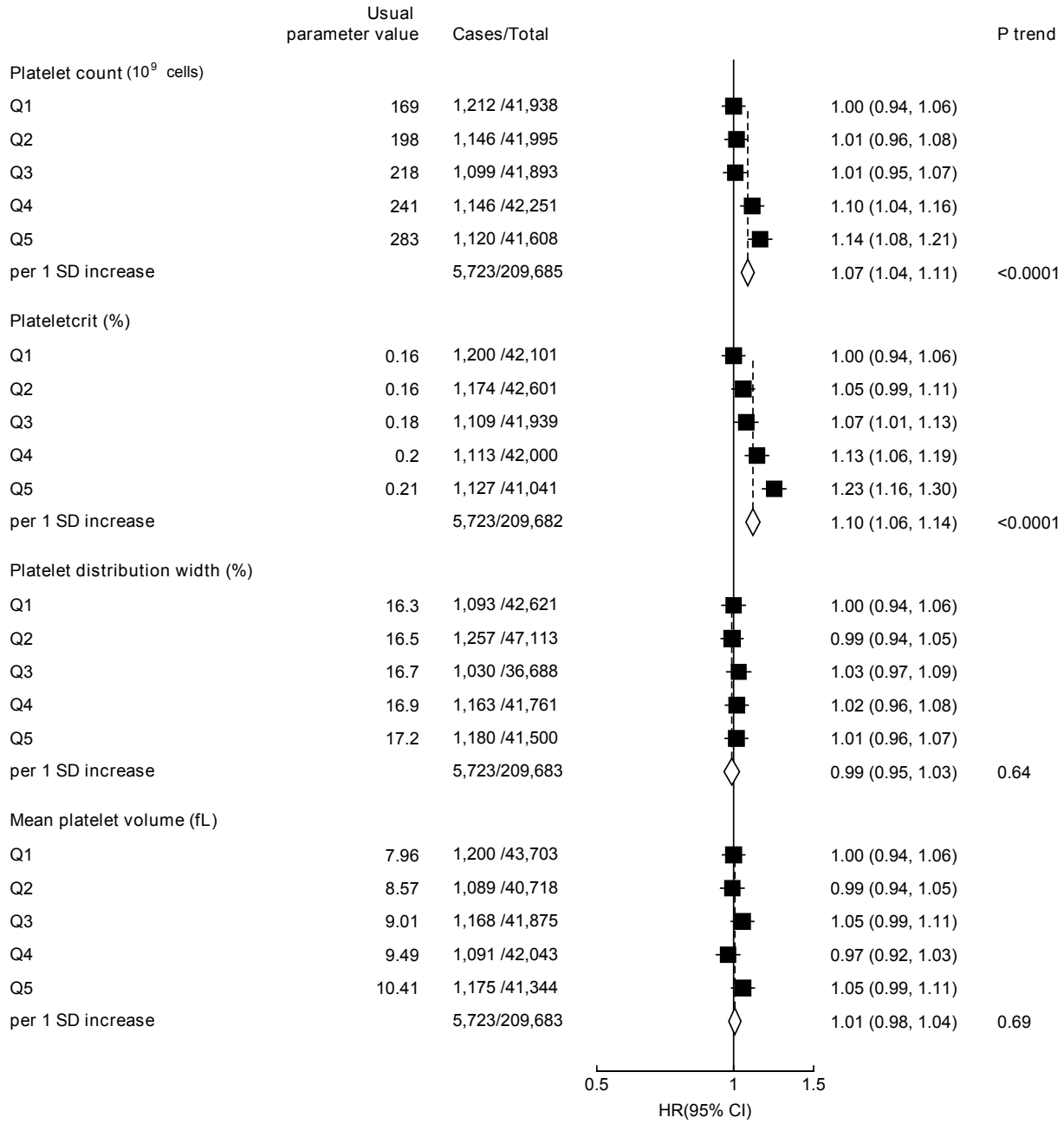


Figure 8.5: Hazard ratio (95% CIs) of prostate cancer diagnosis by fifths of platelet indices.

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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). HRs per 1 SD increase are adjusted for regression dilution bias.

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio; IQR=interquartile range; SD=standard deviation.

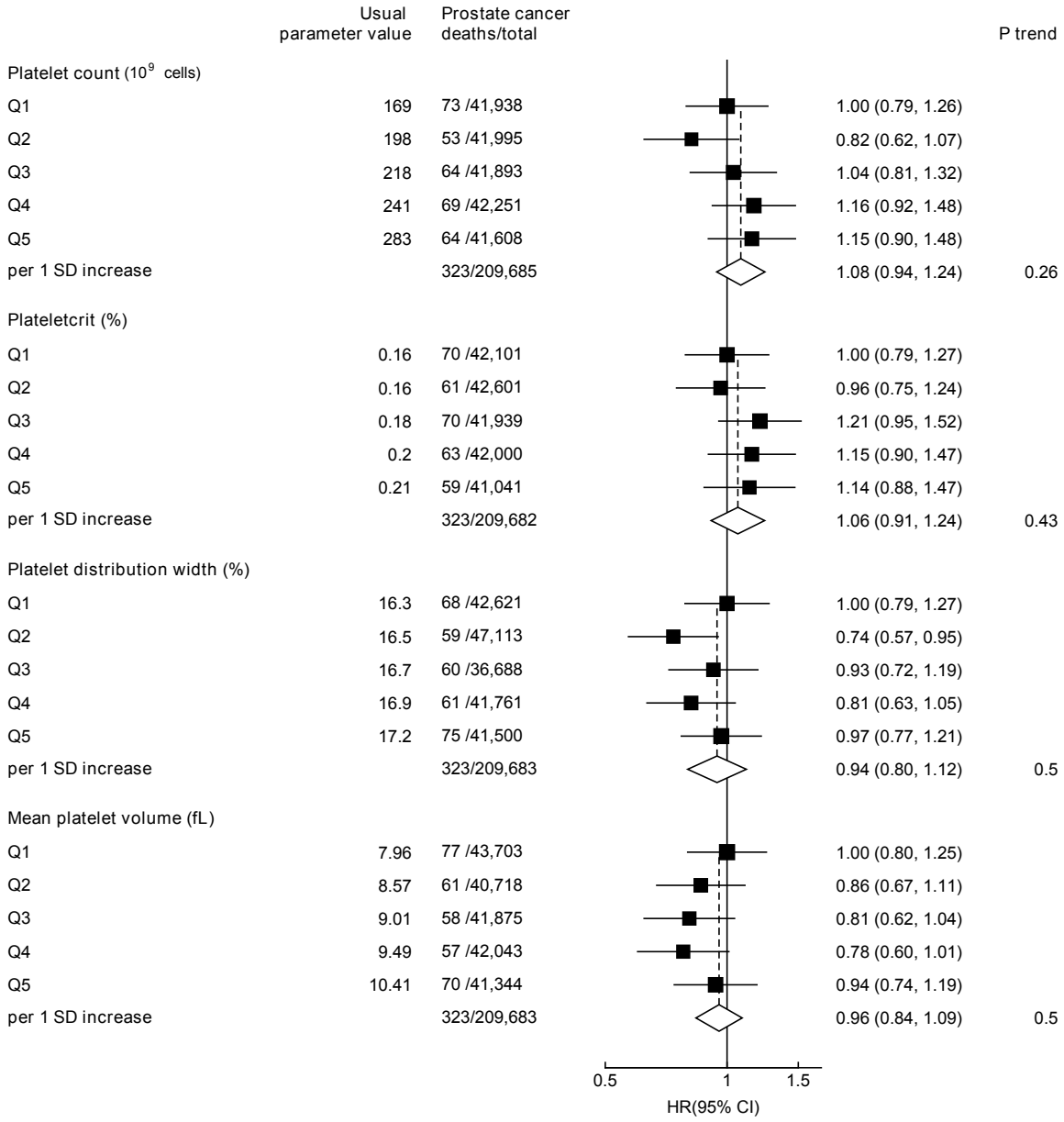


Figure 8.6: Hazard ratio (95% CIs) of prostate cancer mortality by fifths of platelet indices.

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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). HRs per 1 SD increase are adjusted for regression dilution bias.

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio; IQR=interquartile range; SD=standard deviation.

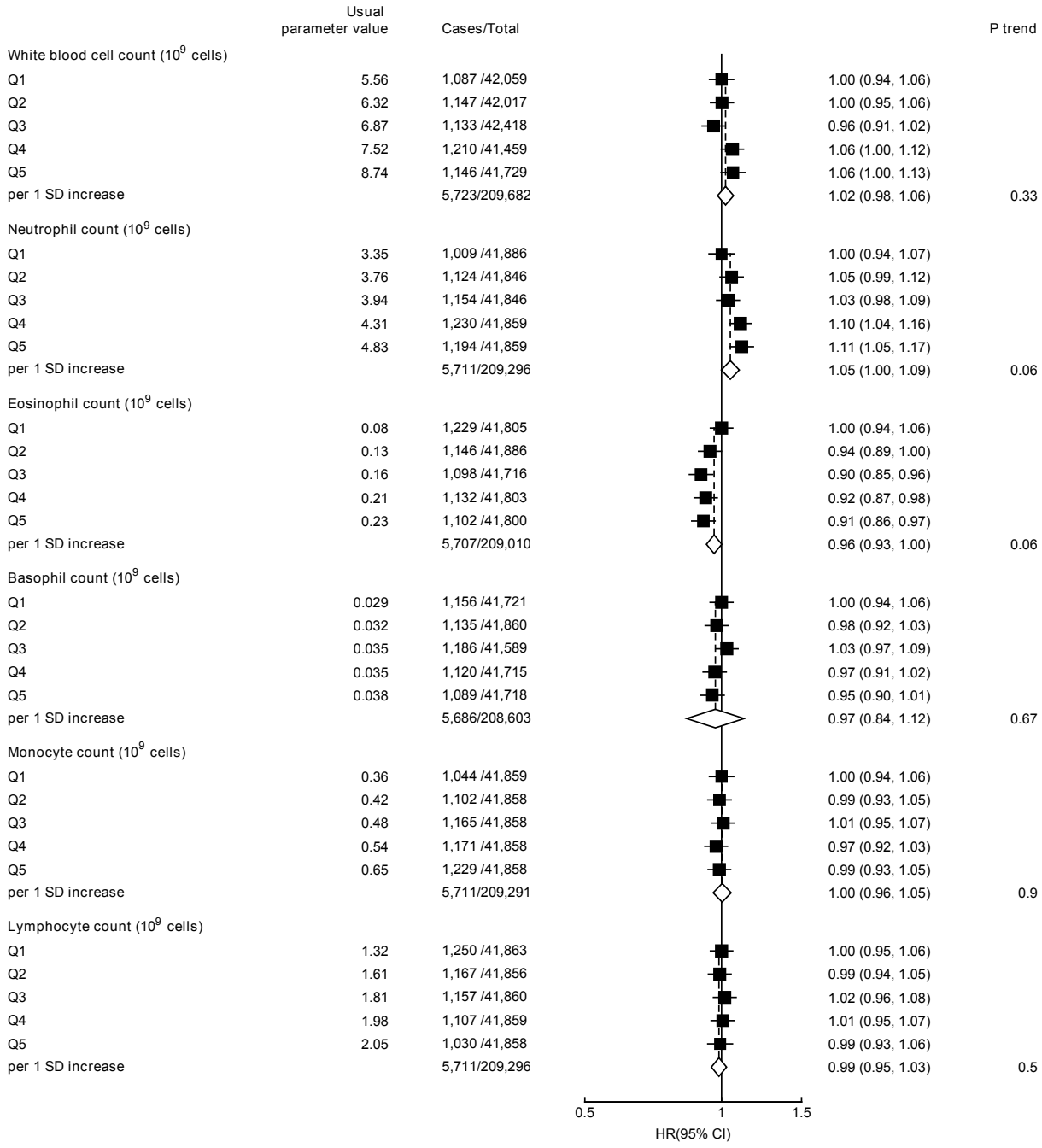


Figure 8.7: Hazard ratio (95% CIs) of prostate cancer diagnosis by fifths of white blood cell indices*.

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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). HRs per 1 SD increase are adjusted for regression dilution bias.

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio; IQR=interquartile range; SD=standard deviation.

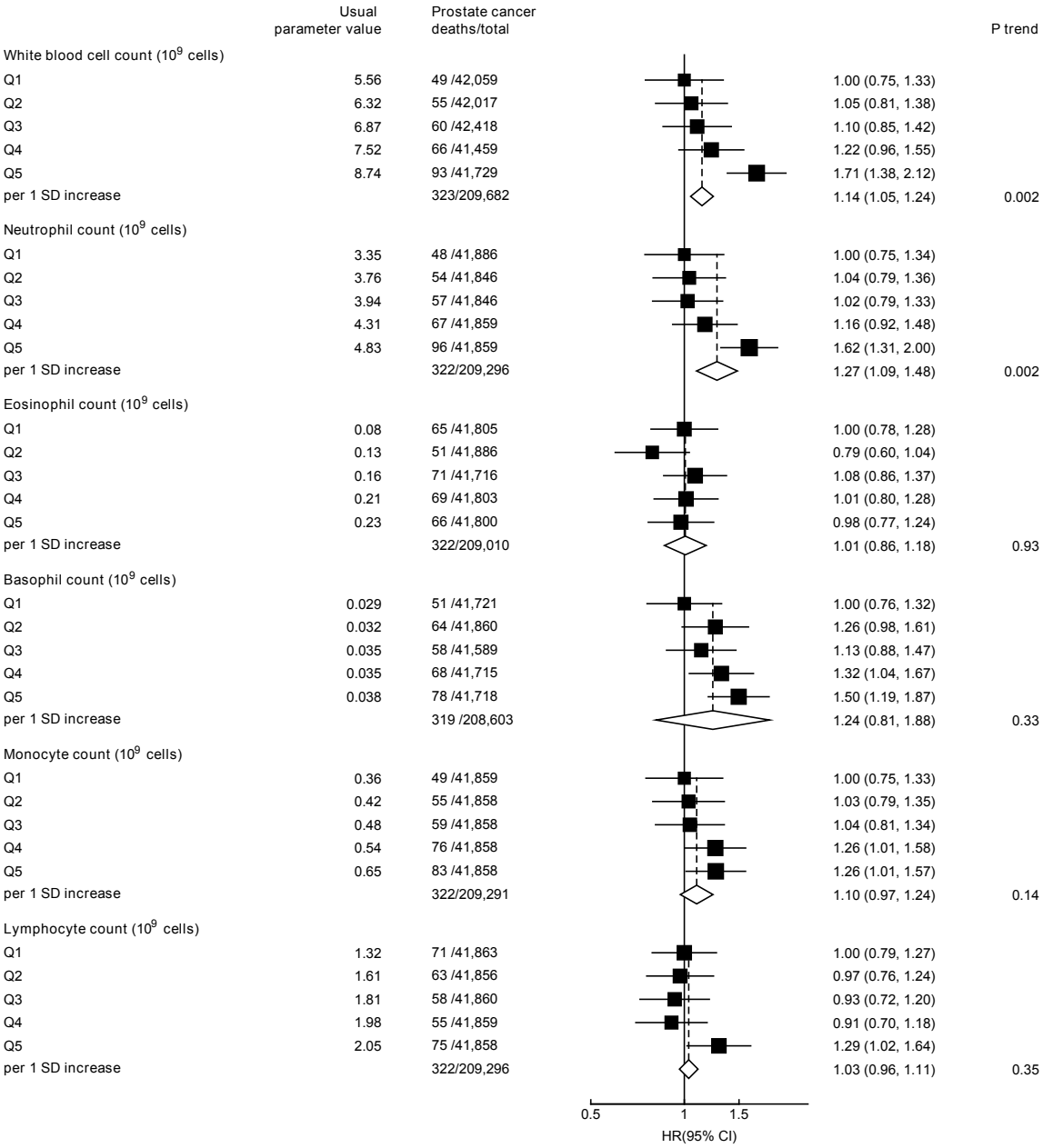


Figure 8.8: Hazard ratio (95% CIs) of prostate cancer mortality by fifths of white blood cell indices*.

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), 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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). HRs per 1 SD increase are adjusted for regression dilution bias.

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio; IQR=interquartile range; SD=standard deviation.

Table 8.5: Hazard ratios for prostate cancer diagnosis and mortality per 1 SD increase (95% CI) in blood parameters with and without correction for regression dilution bias.

	Incidence		Mortality	
	Uncorrected for RDB	Corrected for RDB	Uncorrected for RDB	Corrected for RDB
Red blood cell				
Red blood cell count (10^{12} cells/L)	1.07 (1.04 - 1.10)	1.09 (1.05 - 1.13)	1.06 (0.95 - 1.19)	1.08 (0.94 - 1.25)
Red blood cell distribution width (%)	1.02 (1.00 - 1.05)	1.03 (1.00 - 1.07)	1.00 (0.89 - 1.11)	0.99 (0.85 - 1.17)
Haematocrit (%)	1.02 (1.00 - 1.05)	1.03 (0.99 - 1.08)	1.03 (0.92 - 1.14)	1.04 (0.88 - 1.22)
Haemoglobin concentration (g/dL)	1.01 (0.98 - 1.03)	1.01 (0.97 - 1.05)	1.01 (0.90 - 1.12)	1.01 (0.87 - 1.17)
Mean corpuscular volume (fL)	0.92 (0.90 - 0.95)	0.90 (0.87 - 0.93)	0.94 (0.83 - 1.05)	0.92 (0.79 - 1.06)
Mean corpuscular haemoglobin (pg)	0.92 (0.89 - 0.94)	0.90 (0.87 - 0.93)	0.91 (0.81 - 1.02)	0.89 (0.77 - 1.03)
Mean corpuscular haemoglobin concentration (g/dL)	0.97 (0.94 - 0.99)	0.87 (0.77 - 0.97)	0.94 (0.83 - 1.06)	0.77 (0.46 - 1.27)
Mean sphered cell volume (fL)	0.93 (0.91 - 0.96)	0.91 (0.87 - 0.94)	0.96 (0.85 - 1.07)	0.94 (0.80 - 1.11)
Reticulocyte count (10^{12} cells/L)	0.99 (0.96 - 1.02)	0.99 (0.93 - 1.04)	1.02 (0.95 - 1.10)	1.04 (0.90 - 1.20)
Mean reticulocyte volume (fL)	0.92 (0.90 - 0.95)	0.87 (0.83 - 0.91)	0.95 (0.85 - 1.07)	0.92 (0.76 - 1.12)
Immature reticulocyte fraction	0.98 (0.96 - 1.01)	0.96 (0.90 - 1.03)	0.94 (0.83 - 1.05)	0.86 (0.66 - 1.12)
Platelet				
Platelet count (10^9 cells/L)	1.06 (1.03 - 1.08)	1.07 (1.04 - 1.11)	1.06 (0.96 - 1.18)	1.08 (0.94 - 1.24)
Plateletcrit (%)	1.07 (1.04 - 1.10)	1.10 (1.06 - 1.14)	1.04 (0.94 - 1.16)	1.06 (0.91 - 1.24)
Platelet distribution width (%)	0.99 (0.97 - 1.02)	0.99 (0.95 - 1.03)	0.96 (0.86 - 1.08)	0.94 (0.80 - 1.12)
Mean platelet volume (fL)	1.01 (0.98 - 1.03)	1.01 (0.98 - 1.04)	0.96 (0.86 - 1.07)	0.96 (0.84 - 1.09)
White blood cell				
White blood cell count (10^9 cells/L)	1.01 (0.99 - 1.04)	1.02 (0.98 - 1.06)	1.09 (1.03 - 1.16)	1.14 (1.05 - 1.24)
Neutrophil count (10^9 cells/L)	1.03 (1.00 - 1.06)	1.05 (1.00 - 1.09)	1.16 (1.06 - 1.27)	1.27 (1.09 - 1.48)
Eosinophil count (10^9 cells/L)	0.97 (0.95 - 1.00)	0.96 (0.93 - 1.00)	1.01 (0.90 - 1.12)	1.01 (0.86 - 1.18)
Basophil count (10^9 cells/L)	0.99 (0.96 - 1.02)	0.97 (0.84 - 1.12)	1.04 (0.96 - 1.14)	1.24 (0.81 - 1.88)
Monocyte count (10^9 cells/L)	1.00 (0.97 - 1.03)	1.00 (0.96 - 1.05)	1.06 (0.98 - 1.14)	1.10 (0.97 - 1.24)
Lymphocyte count (10^9 cells/L)	0.99 (0.96 - 1.02)	0.99 (0.95 - 1.03)	1.03 (0.97 - 1.08)	1.03 (0.96 - 1.11)

Hazard ratios per 1 SD increase 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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown). Estimates are corrected for regression dilution bias using the MacMahon-Peto regression dilution ratio.

Abbreviations: BMI=body mass index; CI=confidence interval; HR=hazard ratio; RDB=regression dilution bias; SD=standard deviation.

8.3.3 Further analyses

Subgroup analyses

There was some evidence of heterogeneity in the associations of red blood cell and platelet parameters with prostate cancer risk by time to diagnosis, with stronger associations with prostate cancer risk in men diagnosed ≥5 years from the baseline blood collection (HR per 1 SD increase=1.14, 95% CI 1.09-1.20; P_{hetero}=0.001; HR=0.86, 0.82-0.90, P_{hetero}=0.001; HR=0.87, 0.82-0.91, P_{hetero}=0.005; HR=0.86, 0.82-0.91, P_{hetero}=0.0002; HR=0.85, 0.81-0.89, P_{hetero}=0.0001; HR=1.11, 1.07-1.16, P_{hetero}=0.004; HR=1.13, 1.08-1.18, P_{hetero}=0.002, for red blood cell count, MCV, MCH, MSCV, mean reticulocyte volume, platelet count and plateletcrit parameters. respectively) (Table 8.6). There was no evidence of heterogeneity in the associations with prostate cancer risk by age at diagnosis (Table 8.7) or age at recruitment, smoking, BMI and diabetes status (P_{het}>0.01 for all; Figures 8.9-8.12).

Sensitivity analyses

Further adjustment for family history of prostate cancer and date of blood collection made no appreciable difference to the associations. Associations remained consistent following the exclusion of i) men who reported regularly taking multivitamins/folate supplementation, ii) medications which are known to influence haematological parameters, and iii) following the exclusion of outliers (results not shown).

Table 8.6: Hazard ratios for prostate cancer diagnosis per 1 SD increase in blood parameters by time to diagnosis.

	Diagnosed <5 years from baseline		Diagnosed ≥5 years from baseline		
	Cases	HR (95% CI)	Cases	HR (95% CI)	P_{het}
Red blood cell					
Red blood cell count (10^{12} cells/L)	3826	1.03(1.00-1.07)	1897	1.14(1.09-1.20)	0.001
Red blood cell distribution width (%)	3826	1.04(1.01-1.07)	1897	1.00(0.95-1.04)	0.16
Haematocrit (%)	3826	1.01(0.98-1.04)	1897	1.05(1.01-1.10)	0.11
Haemoglobin concentration (g/dL)	3826	0.99(0.96-1.02)	1897	1.04(0.99-1.09)	0.11
Mean corpuscular volume (fL)	3826	0.95(0.92-0.99)	1897	0.86(0.82-0.90)	0.001
Mean corpuscular haemoglobin (pg)	3826	0.94(0.91-0.98)	1897	0.87(0.82-0.91)	0.005
Mean corpuscular haemoglobin conc (g/dL)	3826	0.96(0.93-1.00)	1897	0.97(0.93-1.02)	0.82
Mean sphered cell volume (fL)	3756	0.97(0.94-1.00)	1876	0.86(0.82-0.91)	<0.001
Reticulocyte count (10^{12} cells/L)	3755	0.99(0.95-1.03)	1876	1.00(0.95-1.04)	0.83
Mean reticulocyte volume (fL)	3755	0.96(0.93-0.99)	1876	0.85(0.81-0.89)	<0.001
Immature reticulocyte fraction	3755	0.99(0.96-1.03)	1876	0.97(0.92-1.02)	0.45
Platelet					
Platelet count (10^9 cells/L)	3826	1.03(1.00-1.06)	1897	1.11(1.07-1.16)	0.004
Plateletcrit (%)	3826	1.04(1.01-1.07)	1897	1.13(1.08-1.18)	0.002
Platelet distribution width (%)	3826	1.00(0.97-1.04)	1897	0.98(0.93-1.02)	0.32
Mean platelet volume (fL)	3758	1.00(0.97-1.03)	1856	1.01(0.97-1.06)	0.76
White blood cell					
White blood cell count (10^9 cells/L)	3817	1.00(0.97-1.04)	1893	1.04(0.99-1.09)	0.27
Neutrophil count (10^9 cells/L)	3815	1.01(0.98-1.05)	1896	1.05(1.00-1.10)	0.21
Eosinophil count (10^9 cells/L)	3813	0.97(0.94-1.01)	1894	0.97(0.93-1.02)	0.98
Basophil count (10^9 cells/L)	3799	1.00(0.96-1.03)	1887	0.99(0.93-1.04)	0.71
Monocyte count (10^9 cells/L)	3815	1.01(0.98-1.04)	1896	0.99(0.94-1.04)	0.49
Lymphocyte count (10^9 cells/L)	3815	0.99(0.95-1.02)	1896	1.00(0.95-1.05)	0.73

HRs per 1 SD increase 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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown).
P_{het} was estimated using stratified Cox models based on competing risks (<5 and ≥5 years from recruitment to diagnosis), and tested using the χ^2 test of heterogeneity.
Abbreviations: BMI=body mass index; CI=confidence interval; HR=hazard ratio; SD=standard deviation.

Table 8.7: Hazard ratios for prostate cancer diagnosis per 1 SD increase in blood parameters by age at diagnosis.

	<65 years at diagnosis		≥65 years at diagnosis		
	Cases	HR (95% CI)	Cases	HR (95% CI)	P_{het}
Red blood cell					
Red blood cell count (10^{12} cells/L)	2049	1.09(1.04-1.14)	3674	1.06(1.03-1.10)	0.42
Red blood cell distribution width (%)	2049	1.03(0.98-1.08)	3674	1.02(0.99-1.05)	0.75
Haematocrit (%)	2049	1.04(0.99-1.09)	3674	1.01(0.98-1.05)	0.39
Haemoglobin concentration (g/dL)	2049	1.02(0.98-1.07)	3674	1.00(0.97-1.03)	0.44
Mean corpuscular volume (fL)	2049	0.92(0.88-0.96)	3674	0.92(0.89-0.96)	0.84
Mean corpuscular haemoglobin (pg)	2049	0.90(0.86-0.95)	3674	0.92(0.89-0.95)	0.53
Mean corpuscular haemoglobin conc (g/dL)	2049	0.96(0.91-1.00)	3674	0.97(0.94-1.01)	0.56
Mean sphered cell volume (fL)	2014	0.94(0.89-0.98)	3618	0.93(0.90-0.96)	0.83
Reticulocyte count (10^{12} cells/L)	2013	0.99(0.94-1.04)	3618	1.00(0.96-1.03)	0.79
Mean reticulocyte volume (fL)	2013	0.91(0.87-0.95)	3618	0.93(0.90-0.96)	0.44
Immature reticulocyte fraction	2013	0.98(0.93-1.03)	3618	0.99(0.95-1.02)	0.79
Platelet					
Platelet count (10^9 cells/L)	2049	1.05(1.01-1.10)	3674	1.06(1.03-1.09)	0.75
Plateletcrit (%)	2049	1.06(1.02-1.11)	3674	1.08(1.04-1.11)	0.71
Platelet distribution width (%)	2049	0.99(0.95-1.03)	3674	1.00(0.96-1.03)	0.78
Mean platelet volume (fL)	2007	1.01(0.96-1.05)	3607	1.00(0.97-1.04)	0.93
White blood cell					
White blood cell count (10^9 cells/L)	2045	1.02(0.97-1.07)	3665	1.01(0.98-1.05)	0.89
Neutrophil count (10^9 cells/L)	2045	1.02(0.98-1.07)	3666	1.03(0.99-1.06)	0.85
Eosinophils (10^9 cells/L)	2044	0.99(0.95-1.04)	3663	0.96(0.94-1.00)	0.35
Basophil count (10^9 cells/L)	2038	1.01(0.96-1.05)	3648	0.99(0.95-1.02)	0.50
Monocyte count (10^9 cells/L)	2045	1.00(0.95-1.05)	3666	1.00(0.97-1.04)	0.45
Lymphocyte count (10^9 cells/L)	2045	1.01(0.96-1.06)	3666	1.01(0.98-1.04)	0.97

HRs per 1 SD increase 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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown).
P_{het} was estimated using stratified Cox models based on competing risks (<65 and ≥65 years at diagnosis), and tested using the χ^2 test of heterogeneity.
Abbreviations: BMI=body mass index; CI=confidence interval; HR=hazard ratio; SD=standard deviation.

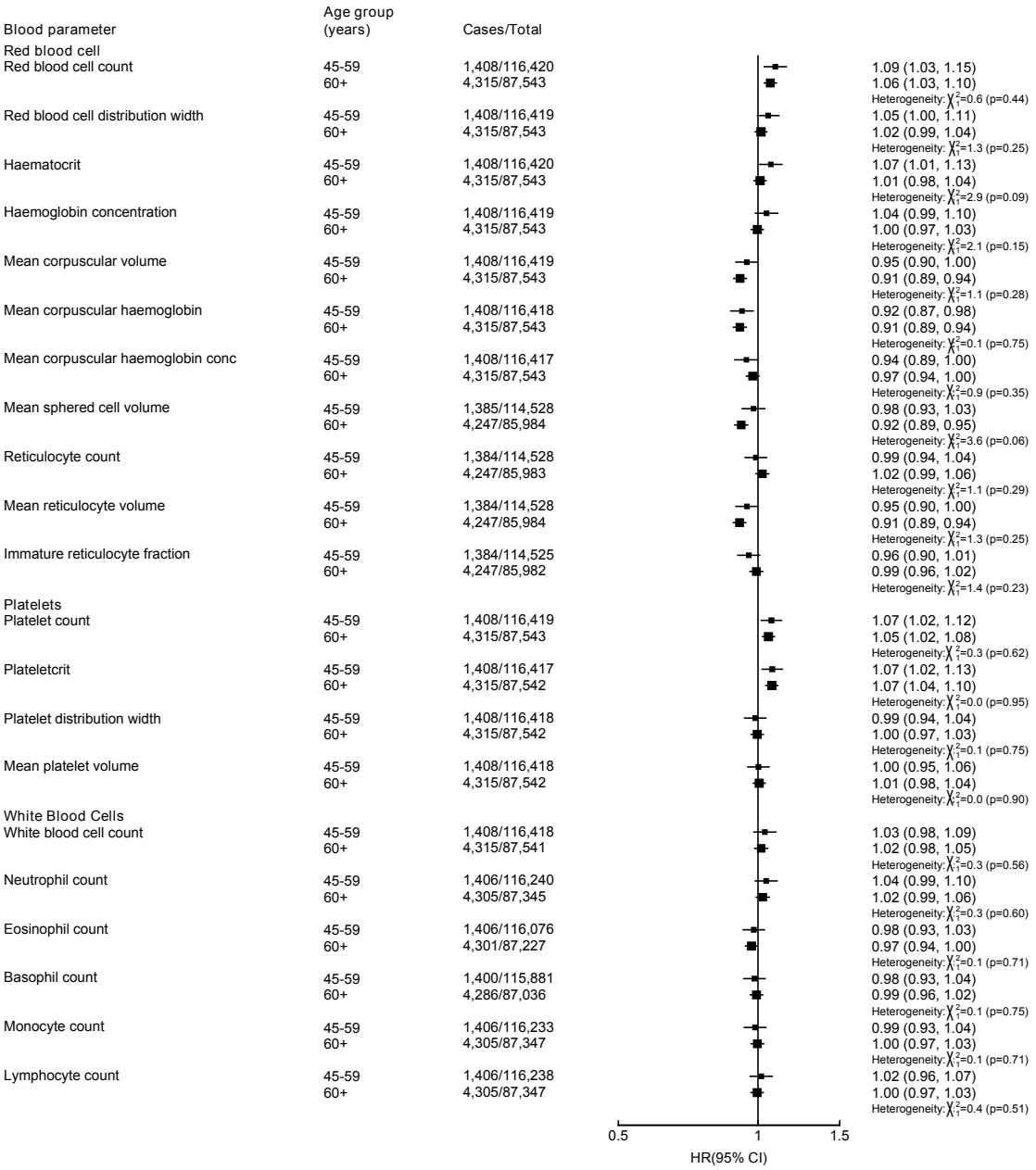


Figure 8.9: Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by age at recruitment.

HRs are stratified by region (10 UK cancer registry regions), 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^2, unknown), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown).

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio.

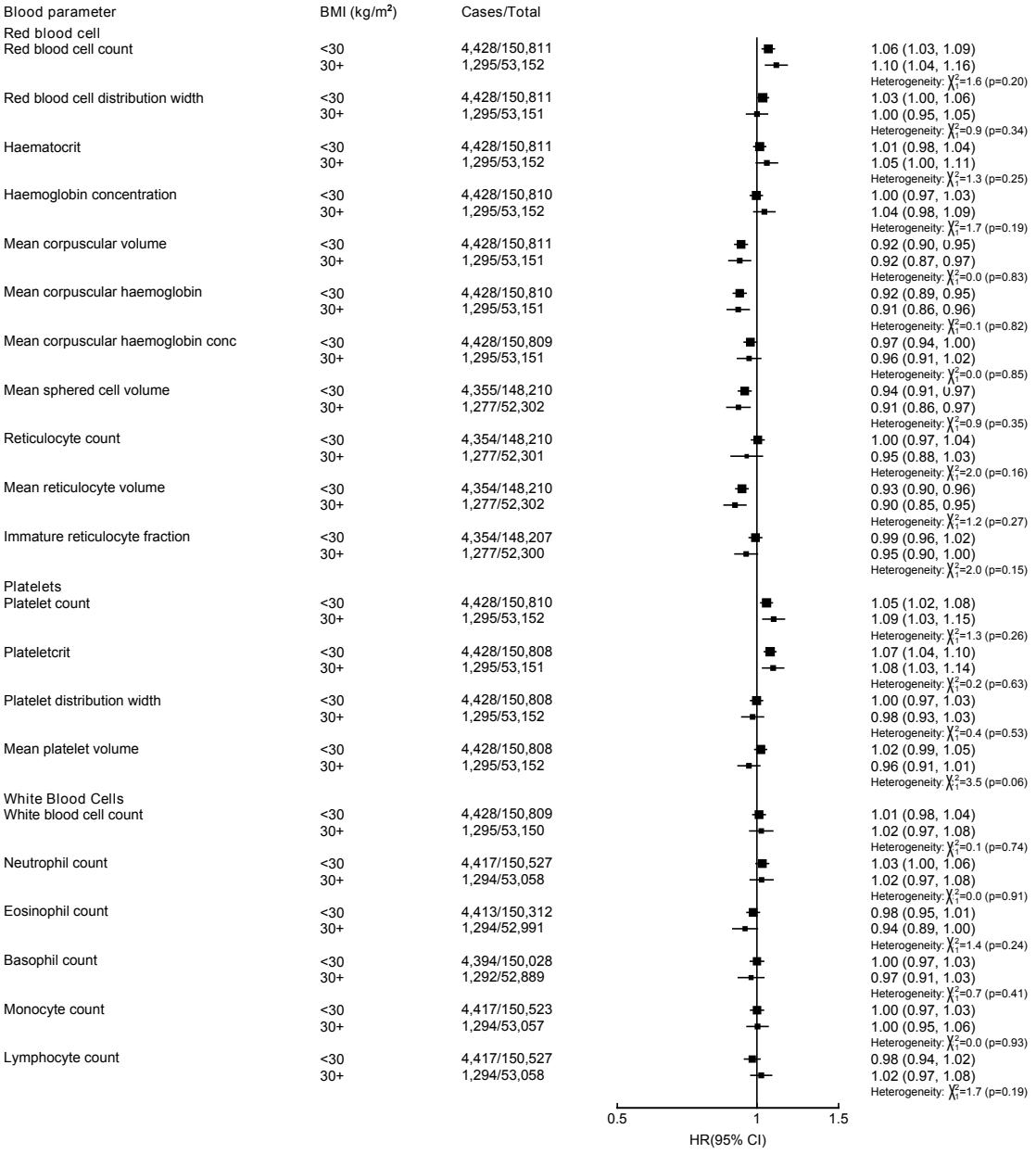


Figure 8.10: Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by category of BMI.

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), 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), smoking status unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown).

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio.

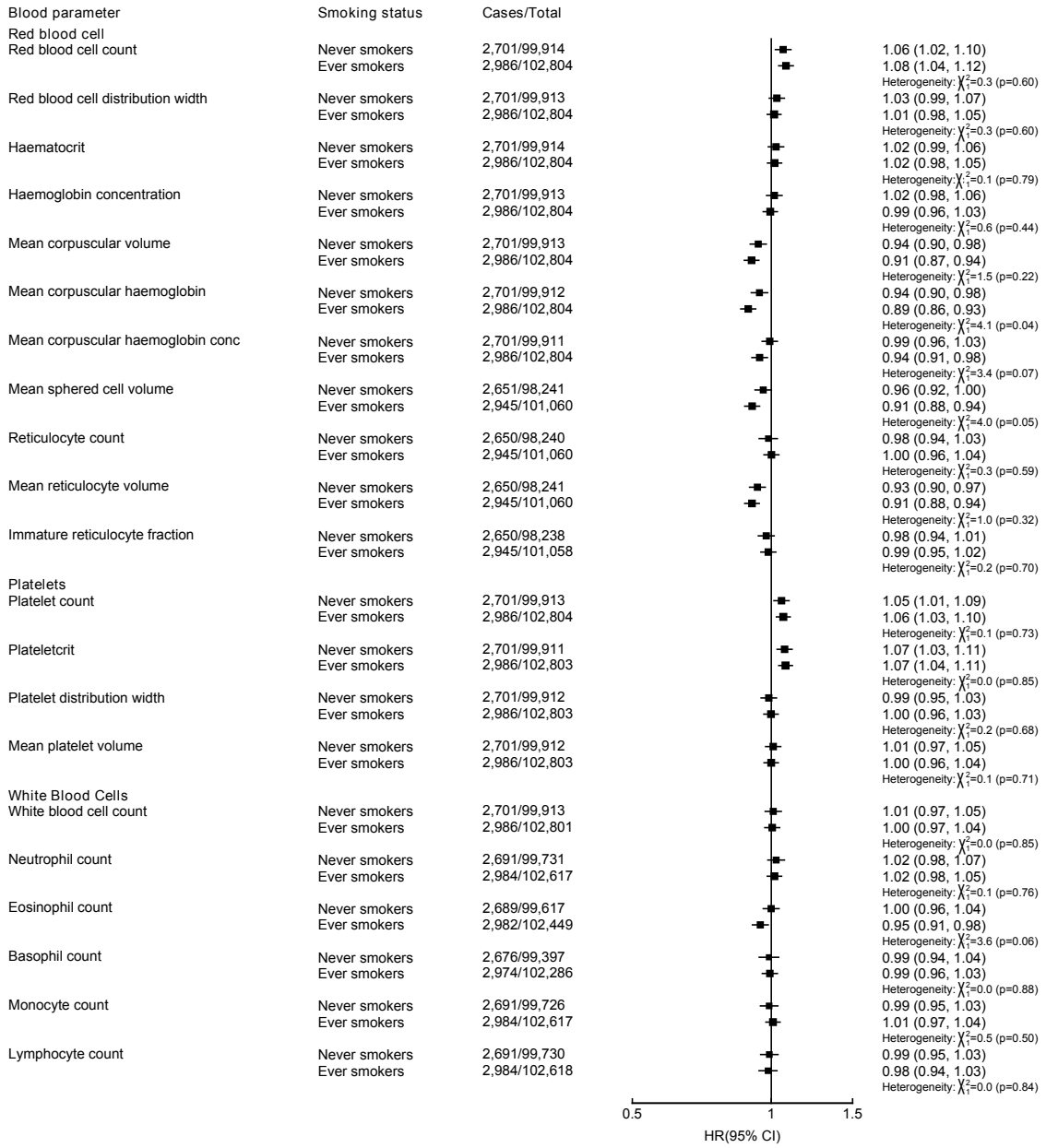


Figure 8.11: Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by smoking status.

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^2, unknown), alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown), and diabetes (no, yes, unknown).

Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio.

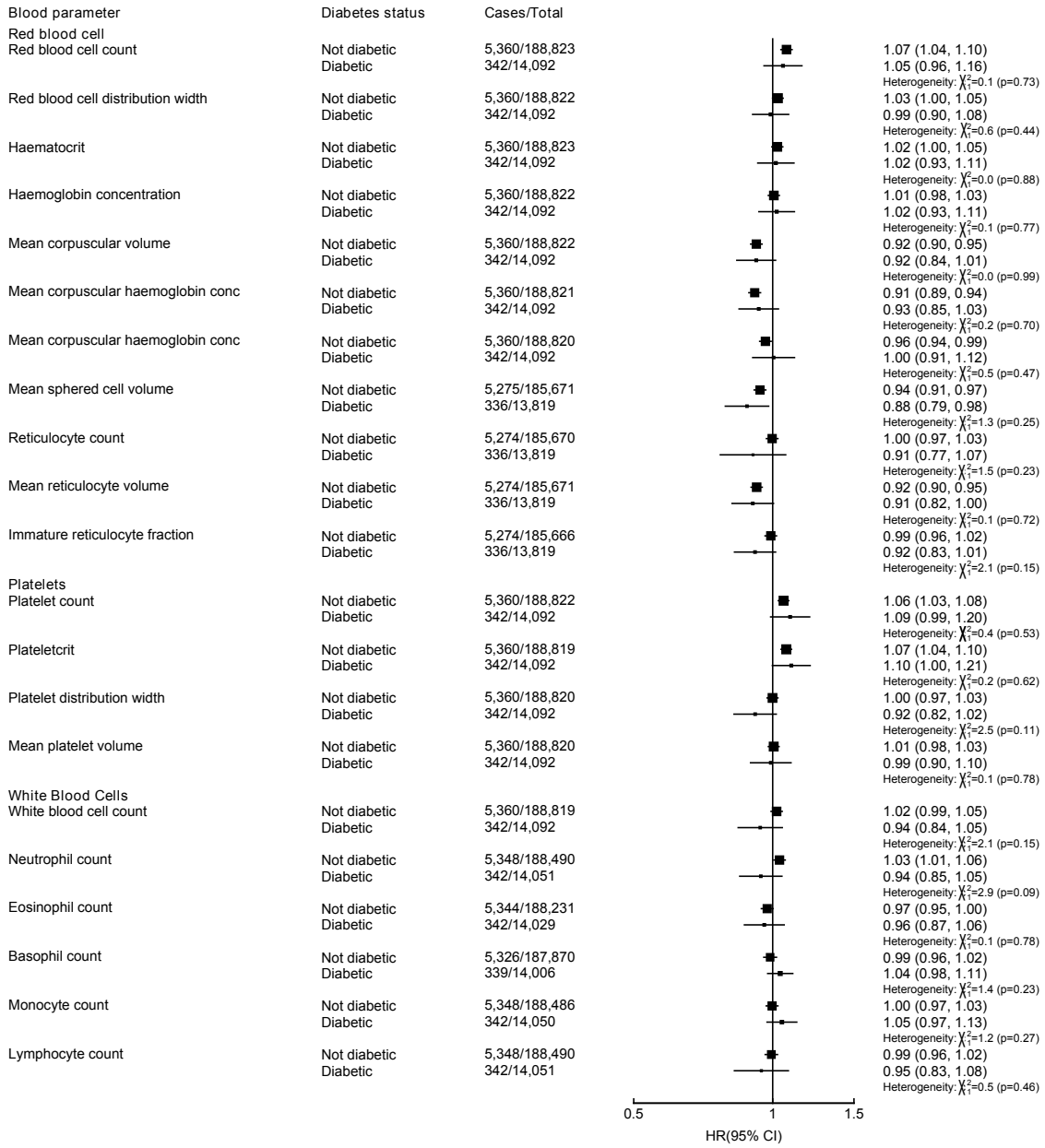


Figure 8.12: Hazard ratio per 1 SD increase of blood parameters with prostate cancer incidence by diabetes status.

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^2, unknown), 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), smoking status unknown) and alcohol consumption (non-drinkers, <1-≤9, ≥10-<20, ≥20 g ethanol/day, unknown).
Abbreviations: BMI=body mass index; CI=confidence intervals; HR=hazard ratio.

8.4 Discussion

Findings from this large prospective study in British men provide evidence for the associations of haematological parameters with prostate cancer risk. Higher red blood cell counts, lower MCV, MCH, MCHC, MSCV and reticulocyte volume, and higher platelet count and plateletcrit were associated with an increased risk of prostate cancer diagnosis, but red blood cell and platelet parameters were not associated with prostate cancer mortality. In contrast, white blood cell counts were not associated with risk of prostate cancer diagnosis, but higher white blood cell and neutrophil counts were associated with an increased risk of prostate cancer mortality.

8.4.1 Associations with red blood cell indices

The observed associations of red blood cell indices with prostate cancer risk may support the hypothesised roles of testosterone, folate, vitamin B_{12} and/or iron in prostate cancer development. Testosterone and dietary factors (folate, vitamin B_{12} and iron) play a role in red blood cell production, and deficiency in any of these factors may cause anaemia [500–505, 522–525]. Previous epidemiological [137, 509, 526] and genetic studies [7, 9, 10, 139, 140, 509] support the possible involvement of these factors in prostate cancer risk (Chapter 2).

8.4.2 Associations with platelet and white blood cell indices

The finding of a positive association between higher plateletcrit and platelet counts and risk of prostate cancer diagnosis, and higher white blood cell and neutrophil counts with increased risk of prostate cancer mortality, may support the hypothesised roles of inflammation and/or infection in prostate cancer development and/or progression [505, 508, 513, 527]. Previous studies have reported that inflammatory markers including leukocyte count are associated with an increased risk of several cancers [515, 516, 518], but the reported associations with prostate cancer risk have been inconsistent [515, 516, 528].

8.4.3 Diet and other factors

Due to the well-established role of nutrients in erythropoiesis, diet may play a role in the observed associations of red blood cell parameters with prostate cancer risk. Previous epidemiological evidence has suggested that folate, vitamin B_{12} and iron may be positively associated with prostate cancer risk [137, 509], but the evidence is inconsistent [139, 510, 511] (Chapter 2).

Red blood cell production is primarily controlled by the kidneys (via erythropoietin) [498, 529]. Previous epidemiological studies have observed a reduced risk of prostate cancer in men with chronic kidney disease [530] therefore, the associations of red blood cell indices with prostate cancer risk may also be related to the correlates of kidney function (including testosterone [531, 532], insulin-like growth factor-I [533] and metabolic syndrome [534]).

8.4.4 Reverse causation

It is possible that the associations observed in the current study may be at least partially explained by reverse causation. Cancer can cause anaemia by producing cytokines which lead to iron sequestration [535], and tumours which infiltrate the bone marrow can suppress haematopoiesis [535]. Associations of red blood cell and platelet indices with prostate cancer diagnosis were stronger with a longer follow-up time, which may suggest that these associations cannot be explained by reverse causation.

Tumours increase platelet indices even during early cancer [536] and may also increase the half-life of neutrophils which can promote tumour growth and metastasis [537, 538]. In this analysis, there was insufficient statistical power to test for heterogeneity in the associations of white blood cell and neutrophil counts with prostate cancer mortality by length of follow-up. Therefore, it is not possible to rule-out reverse causation in the associations of white blood cell counts with prostate cancer mortality.

8.4.5 Strengths and limitations

With data on over 200,000 men, this is the largest prospective study to examine a wide range of haematological measures in relation to prostate cancer specific risk. Blood indices were measured across the entire cohort via standardised methods and participants were well-characterised. The availability of repeat measures in a subsample of men allowed for correction for regression dilution bias, and therefore more accurate estimates of risk.

A limitation of this analysis was that it is not possible to rule-out detection bias. These blood indices may be associated with health status (for instance, low red blood cell count is associated with hypertension, hypothyroidism and congestive heart failure [539], while high white blood cell count is associated with increased total and cardiovascular mortality [540, 541]). Co-morbidities and poor health may affect PSA test attendance [44], but PSA testing attendance after baseline was not known. Stage and grade data are not currently available in UK Biobank, which may inform further the possible role of detection bias by focusing on more clinically aggressive tumour subtypes.

UK Biobank is a predominantly white study population; therefore, it was not possible to test for heterogeneity in the associations by ethnic group. Although the blood measurements were generally consistent between baseline measurement and the repeat measurement, MCHC and basophil count had low regression dilution ratios, which increased the uncertainty of the risk estimates for these indices.

Blood indices were used as surrogate markers of a range of possible risk factors. Circulating folate, vitamin B_{12}, iron and testosterone concentrations or markers of inflammation (such as C-reactive protein) were not available at the time of analysis, therefore it was not possible to investigate further any potential mechanistic factors which may account for the observed associations. Further research using genetic methods such as Mendelian randomization may be beneficial in identifying possible causal associations of circulating hormones, nutritional factors and the immune response [542] with prostate cancer risk.

8.4.6 Conclusion

In conclusion, this analysis of more than 200,000 men observed associations of several haematological parameters and prostate cancer risk. The associations with red blood cell indices with prostate cancer diagnosis are consistent with the hypothesis that testosterone, folate, vitamin B_{12} and iron may be involved in prostate cancer. The positive associations of platelet indices with prostate cancer diagnosis also suggests a possible role for inflammation/infection in prostate cancer development. Higher white blood cell and neutrophil counts were associated with an increased risk of prostate cancer mortality. However, it was not possible to rule-out reverse causation as an explanation for these associations with prostate cancer mortality. Future analyses, such as Mendelian randomization may be beneficial in identifying the possible causal components which may account for these associations.

9

Summary and recommendations for future research

Contents

9.1 Introduction 241
9.2 Summary of findings 241
9.2.1 Findings across analyses and contribution to the literature 243
9.3 Methodological considerations 246
9.3.1 The role of chance 246
9.3.2 The role of bias 247
9.3.3 The role of confounding 254
9.3.4 Summary of limitations 254
9.4 Priorities for future research 256
9.4.1 Hormones in healthy and malignant prostate tissue . . . 256
9.4.2 Triangulation 256
9.4.3 Mediation analysis 259
9.4.4 Aggressive tumour subtypes 259
9.4.5 Collaboration and data pooling 260
9.4.6 Implications for prostate cancer prevention 260
9.5 Conclusion 261

9.1 Introduction

The aim of the research described in this thesis was to investigate the epidemiology and associations of circulating testosterone, insulin-like growth factor-I (IGF-I) and other blood characteristics with prostate cancer risk, by examining:

1. The cross-sectional associations of anthropometric, sociodemographic, and behavioural factors with circulating sex hormones, IGFs and insulin-like growth factor binding proteins (IGFBPs).

2. The associations of circulating testosterone and IGF-I concentrations and clinical surrogates with prostate cancer risk.

3. The associations of complete blood count parameters with prostate cancer diagnosis and mortality, to further examine the role of testosterone and other potential risk factors for prostate cancer.

9.2 Summary of findings

- **Chapter 4:** Used data from the Endogenous Hormones, Nutritional Biomarkers and Prostate Cancer Collaborative Group (EHNBPCCG), an international consortium of prospective studies and prostate cancer risk, to examine the cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating sex hormone concentrations in men. These analyses showed that age and body mass index (BMI) and, to a lesser extent, smoking status and alcohol consumption were associated with circulating sex hormones.

- **Chapter 5:** Investigated the cross-sectional associations of anthropometric, sociodemographic and behavioural factors with circulating IGFs and IGFBPs in men using the EHNBPCCG dataset. Age, BMI, height and ethnic/racial group were associated with circulating IGF and IGFBP concentrations (particularly IGFBP-1 and IGFBP-2). There was some evidence of an association between

smoking status and alcohol consumption and IGFBP concentrations, but associations were generally of a smaller magnitude. Partial correlations between circulating sex hormones and IGFs were also examined, but only modest correlations were observed.

- **Chapter 6:** Investigated a possible threshold association between low free testosterone and prostate cancer risk using the EHNBPCCG dataset. Very low circulating free testosterone concentrations were associated with a reduced risk of prostate cancer. However, there was some suggestion that men with very low free testosterone have an increased risk of high-grade prostate cancer.

- **Chapter 7:** Investigated whether the associations observed in population-based prospective studies of circulating IGF-I and testosterone with prostate cancer risk are consistent in men with clinically extreme concentrations of these hormones, using national hospital admission and mortality records. Men diagnosed with acromegaly (which is characterised by high IGF-I) had an elevated risk of prostate cancer diagnosis and possibly higher prostate cancer mortality. Diagnosis of Klinefelter's syndrome, testicular hypofunction or hypopituitarism was used as a surrogate marker of low testosterone. Men diagnosed with Klinefelter's syndrome had a lower risk of prostate cancer diagnosis. However, there was no evidence of an association between testicular hypofunction or hypopituitarism diagnosis and subsequent prostate cancer diagnosis. Men diagnosed with these hypogonadal-associated diseases had a lower risk of prostate cancer mortality.

- **Chapter 8:** Red blood cells, platelets and white blood cell indices are markers of a range of exposures which may relate to prostate cancer risk, including circulating testosterone, folate, vitmain B_{12}, iron, inflammation and infection. In the UK Biobank, complete blood count and reticulocyte parameters were measured across the entire cohort at study baseline. Higher red blood cell and platelet counts, and lower measures of red blood cell volumes were associated with an increased risk of prostate cancer diagnosis. Higher white blood cell

and neutrophil counts were associated with an increased risk of prostate cancer mortality.

9.2.1 Findings across analyses and contribution to the literature

- **Chapters 4 and 5:** The EHNBPCCG pooled dataset contains the largest collection of endogenous hormone concentrations and prostate cancer status currently available. The cross-sectional analyses in Chapters 4 and 5 increase the evidence base of the known associations of a range of anthropometric, sociodemographic and behavioural factors with circulating sex hormones, IGFs and their binding proteins [297–301, 366, 368].

  The large size of this dataset also enabled the investigation of a number of hormones and binding proteins only previously examined in a few, mostly small individual studies (such as androstanediol glucuronide (A-diol-G), estrone and IGFBP-1 and IGFBP-2) and observed associations which have not previously been reported (including associations of IGFBPs and androstenedione with smoking and the associations of androstenedione, A-diol-G, IGF-II and IGFBP-2 with alcohol). If the observed cross-sectional associations with these less well studied hormones and binding proteins are supported in further analyses, then these results will enhance our understanding of how these factors may affect prostate cancer risk via circulating hormone concentrations.

- **Chapter 6:** Despite evidence from animal models [4, 12, 211], genetic studies [6, 7], and two randomised controlled trials which support a role of androgens in prostate cancer development [8, 9], until now there has been no consistent prospective epidemiological evidence of an association between circulating testosterone concentrations and prostate cancer risk [13]. This discrepancy between experimental and clinical evidence and prospective epidemiological evidence has been under consideration for decades [543]. It has been theorised that this may be due to a threshold association between free testosterone and

prostate cancer risk [19]. Therefore, individual epidemiological studies may have been underpowered to detect this non-linear association.

The finding in Chapter 6, based on a large international consortium, that low circulating free testosterone concentrations are associated with a reduction in overall prostate cancer risk therefore represents a substantial contribution to the literature. Furthermore, the evidence of heterogeneity in the associations by tumour grade, with a possible increased risk for high-grade disease, is also consistent with findings for high-grade prostate cancer reported by randomised controlled trials which aimed to reduce prostate cancer risk by reducing intra-prostatic androgen concentrations [8, 9].

- **Chapter 7:** Previous prospective studies have reported associations of circulating IGF-I [14] and testosterone (Chapter 6) with prostate cancer risk. This chapter tested if associations persisted in the context of clinical extremes, which population-based cohort studies are generally underpowered to investigate.

  Previous analyses have found that men diagnosed with acromegaly have an increased risk of overall cancer diagnosis [459], but no study has had the statistical power to robustly investigate the association between acromegaly and prostate cancer risk. In this analysis, the observed increased risk of prostate cancer diagnosis in men with clinically high IGF-I (acromegaly) is consistent with previous epidemiological evidence that men with higher circulating IGF-I have an increased risk of prostate cancer [14].

  Previous observational studies have found some evidence that Klinefelter's syndrome may be associated with a lower prostate cancer risk, but the evidence was inconclusive [463, 469, 470], and it is not known if this association is consistent across other hypogonadal-associated diseases. In this analysis, the observed reduction in prostate cancer risk in men diagnosed with Klinefelter's syndrome may support the findings in Chapter 6. However, there was no evidence of an association with prostate cancer risk in men diagnosed with testicular hypofunction or hypopituitarism, although these men had a lower

risk of prostate cancer mortality. The interpretation of these results in the context of Chapter 6 is complex, due to the recommendation of prostate cancer screening in men who are prescribed testosterone replacement therapy, and the lack of data on tumour characterisation.

- **Chapter 8:** Previous studies have implicated factors including testosterone, nutrition, inflammation and infection in prostate cancer risk [119, 137]. While several prospective studies have examined the associations between a limited number of blood parameters and risk of cancer at all sites [514–518], few have had the power to describe these exposures in relation to prostate cancer risk [514, 515]. The analyses reported in this thesis are unique due to the breadth of blood parameters examined, the large sample size with blood collected across the whole cohort, and the availability of repeated measurement data. The associations with red blood cell parameters with prostate cancer risk are consistent with the possible involvement of testosterone in prostate cancer risk reported in Chapter 6. The associations of platelet indices with prostate cancer incidence and white blood cells with prostate cancer mortality also support the possible role of inflammation and/or infection in prostate cancer development. However, other correlates of blood indices are not well understood and data on other blood biomarkers were not available at the time of analysis. Therefore, it was not possible to infer further the possible underlying driving mechanisms of the observed associations of these haematological indices with prostate cancer risk.

9.3 Methodological considerations

The work presented in this thesis has several strengths, particularly the large size of the datasets analysed and the diversity of the study designs. However, the analyses have a number of limitations which are described below.

9.3.1 The role of chance

Multiple testing

Because of the breadth of hormone and participant characterisation data available, multiple tests were carried out in each analysis. A greater number of statistical tests increases the probability of type I errors. This particularly may be the case in Chapters 4, 5 and 8, where the statistical threshold was lowered to $p<0.01$ to reduce the probability of these errors. Because the analyses were exploratory in these chapters, the analyses did not adjust further for multiple testing to avoid rejecting potentially important findings [307]. However, the associations observed, particularly those with weaker p-values, have been reported in the context of evidence from other studies.

Sample size and statistical power

This thesis analysed some of the largest datasets available worldwide. This increased the power to detect associations with overall prostate cancer risk. However, there was still relatively limited power to examine associations with the small number of more clinically aggressive prostate tumour subtypes (Chapters 6-8) and rare hormone-related diseases (Chapter 7). Therefore, still larger sample sizes (which may be achieved by incorporating more cohort studies including UK Biobank into the EHNBPCCG dataset) and longer follow-up may be required to robustly estimate possible associations with these aggressive prostate tumour subtypes.

9.3.2 The role of bias

Selection bias

Studies in these analyses generally contain healthier participants with a higher sociodemographic status than the underlying population [519, 544, 545]. Only ∼6% of the eligible population who were contacted enrolled in UK Biobank [279] and this study population also has a greater prevalence of diagnosed prostate cancer than the underlying population, which is probably due to a higher PSA test attendance [519]. If the exposure of interest and the outcome are both associated with the probability of being in the study population, being a participant of the study population may introduce collider bias to the analyses, which may lead to spurious associations between otherwise independent variables [546, 547]. Recent research indicates that risk factor-endpoint associations are directionally consistent between the UK Biobank and nationally-sampled British cohort studies with higher response rates (mean 68%) [548],

Generalisability

The lack of racial/ethnic diversity in some datasets was another limitation, although this homogeneity may enable more robust estimates of risk for participants within this group. The EHNBPCCG contained three studies based on Japanese populations (the Japan Collaborative Cohort Study, Japan Public Health Center-based prospective study and Japan-Hawaii Cancer Study), and the Multiethnic Cohort aimed to recruit an ethnically diverse population. Analyses in Chapter 6 on free testosterone concentrations and prostate cancer risk indicated that associations were generally consistent across these studies, but it was not possible to verify this for the exposures investigated in Chapters 7 and 8.

Measurement error

Measurement error refers to the deviation of a measurement from the "true" value [549]. Non-differential errors will reduce the precision of the risk estimates and

attenuate associations towards the null, while differential errors between cases and controls can bias the risk estimates in either direction [550].

Intra-study measurements: Variation in blood sample collection and storage, assay quality, contamination, isolation of biomarkers, freeze-thaw cycles of serum samples and blood clotting will usually result in random measurement error (assuming that samples for men who do and do not develop prostate cancer are handled similarly) [550].

Circulating free testosterone concentrations were not directly measured in any study. Therefore, concentrations were estimated using the law of mass action [302], which is based on testosterone and sex hormone binding globulin (SHBG) concentrations and assumes a constant albumin concentration. Although this is a commonly used method for estimating free testosterone concentrations, this was not validated within each individual study via equilibrium dialysis [159].

ICD-9 and 10 codes were used to identify and "follow-up" exposure and reference cohorts in Chapter 7. Coding errors may have led to some misclassification of cohorts and prostate cancer status. This may lead to an underestimation of risk, but any misclassification is likely to be small [497].

Inter-study measurements: Sample handling, assay method and the likelihood of measurement error may vary by study. There were large differences in hormone concentrations between studies that contributed data to the EHNBPCCG. These differences may be due to absolute differences in hormones between study populations or be explained by differences in blood collection, storage and assays [428]. Therefore, study-specific tenths were derived to account for these differences (Chapter 6).

The UK Biobank used a standardised method of blood collection, storage and measurement across the entire cohort [280, 281]. This standardised approach makes it possible to investigate associations of biomarkers with risk using cut-points defined by the distribution across the whole cohort.

Some studies had long recruitment periods and so measurements were not carried out within a single assay batch. Therefore, laboratory drift may bias results.

Therefore, in the EHNBPCCG analysis, cases and controls were matched by batch or date of blood collection, where applicable (Chapter 6), while in UK Biobank year of blood collection was adjusted for as an additional sensitivity analysis (Chapter 7).

Regression dilution bias: All analyses using the EHNBPCCG were based on single measurements, on the assumption that these measures are representative of participants' hormone concentrations over the medium-to-long term. While IGF-I and testosterone have moderately good reproducibility over a one-to-five year period [355, 368, 380, 551], the reproducibility of these measures over the longer term is unknown. Random measurement error may also lead to an attenuation of the recorded associations (i.e. regression dilution bias [290]).

UK Biobank had repeat measurements in a subset of men, therefore it was possible to adjust for regression dilution bias [290]. Although the reproducibility of many of the blood measurements were generally good between baseline and repeat assessment, mean corpuscular haemoglobin concentration and basophil count had low regression dilution ratios, which increased the uncertainty of these risk estimates after adjustment for regression dilution bias.

The free hormone hypothesis: This thesis assumed that bound testosterone is unable to pass out of the bloodstream into the prostate tissue [157]. Measurements of hormone concentrations in tissue samples indicate that circulating total testosterone is a poor surrogate marker of intra-prostatic testosterone [552]. The free hormone hypothesis is the most prominent theory of testosterone's biological activity and predicts that circulating low free testosterone, rather than total testosterone, should more closely reflect tissue concentrations. However, the extent to which endogenous circulating free testosterone concentrations are suitable surrogate markers of intra-prostatic androgen activity has not been directly validated.

Surrogate indicators of hormones and other factors: In Chapters 7 and 8, hormones were not measured, instead clinical diseases or blood indices were used as surrogate measures of hormones and/or other factors (Chapters 7 & 8). Although these surrogate markers were chosen *a priori* based on pre-existing evidence of associations with the exposures of interest (testosterone, IGF-I, nutrients,

inflammation/infection), there is likely to be some variation in these exposures between men due to factors such as disease treatment, unusual disease presentations or the other determinants of these surrogates.

Pre-analytical measurement error: Some hormones are affected by factors such as diurnal rhythm or fasting status. This may lead to some measurement error if these factors are not accounted for. This error is likely to be non-differential. These factors were conditioned on the matching of cases to controls (Chapter 6) or adjusted for (Chapters 4 & 5), where available.

Other baseline measurements and missing data

To collect data for the EHNBPCCG dataset, potential principal investigators were contacted to request data. Only limited data were requested, therefore the baseline participant characterisation information available in this dataset may be less detailed than in the original study. Measures such as physical activity, nutrition or co-morbidities were not requested. Furthermore, many studies did not collect data such as smoking status, time of blood collection or alcohol consumption.

The UK Biobank cohort contained comparatively more complete and comprehensive data on the characteristics of participants (Chapter 8) in comparison with the national hospital and death records (Chapter 7). Data not collected and missing data limited the ability to adjust for potential confounders. Data such as family history of prostate cancer and BMI were not missing at random therefore multiple imputation was not considered to be appropriate [553].

Data from these studies were generally collected at cohort entry, therefore it is likely that some of these lifestyle and lifestyle-related factors such as BMI, smoking and alcohol will have changed during follow-up, which may attenuate the associations of these potential confounders with prostate cancer risk.

Detection bias

Given the high incidence of latent prostate tumours and the widespread use of prostate-specific antigen (PSA) testing [40], it is important to consider the possibility of detection bias in the investigation of potential prostate cancer risk factors.

Socioeconomic status and health: It is well established that men who have higher levels of education and are more health conscious are more likely to undergo PSA testing [43, 44], and therefore are more likely to be diagnosed with prostate cancer.

The cross-sectional analyses showed only modest associations of factors such as education with circulating sex hormones and IGFs, but larger magnitudes of the associations were observed with BMI (Chapters 4 and 5). However, men diagnosed with Klinefelter's syndrome had higher levels of socioeconomic deprivation than the reference cohort (Chapter 7). In the UK Biobank population, there was variation in blood count indices by factors such BMI, ethnicity and physical activity. Where possible the analyses adjusted for these factors and tested for heterogeneity in the associations by these factors.

Co-morbidities may play a role in the associations with prostate cancer. Hypogonadism is associated with poor health [554]. In the case of the haematological parameters, low red blood cell counts are associated with an increased risk of hypertension, hypothyroidism and congestive heart failure [539]. Men with poorer health may have a lower PSA test attendance [44]. Therefore, the observed associations of low free testosterone and haematological indices with prostate cancer risk may be partially explained by detection bias. In contrast, men diagnosed with clinical hypogonadism and receiving testosterone replacement therapy are recommended to undergo regular prostate cancer screening, which is likely to increase prostate tumour detection [481].

Hormones, PSA and prostate morphology: Hormones, particularly testosterone, have an important role in normal prostate function, and influence the size and morphology of the prostate gland [183]. Androgen receptor signalling also directly controls PSA production [194, 204, 445] and men with very low circulating free testosterone had lower PSA concentrations (Chapter 6). Therefore, circulating hormone concentrations may affect prostate tissue or PSA concentrations in such a way which that may bias the likelihood of prostate diagnosis and possibly the diagnosed tumour grade rather than directly affecting prostate tumour

development/progression [555]. This has been a source of on-going debate in the randomised controlled trials of 5-α reductase inhibitors and prostate cancer [443, 444, 556].

Study designs

This thesis used a range of different study designs and datasets. Investigations were based on cross-sectional (Chapters 4 & 5), retrospective (Chapter 7) and prospective analyses (Chapters 6 & 8). The analysis of national hospital and mortality records is technically considered a retrospective cohort design because the exposures and cases are identified post-hoc (Chapter 7). However, exposures were identified using national records, rather than participant recall, therefore this study shares many of the advantages of conventional prospective cohort studies [557].

Analyses in the EHNBPCCG pooled dataset are advantageous over a traditional meta-analysis as they are based on both published and unpublished data and therefore may be less prone to publication bias. This large pooled dataset with individual participant data also allows for greater flexibility of analyses.

No analyses were based on an intervention, therefore it is only possible to infer an association, rather than causation [558].

Participant follow-up and tumour characterisation

Loss to follow-up will reduce statistical power and introduce bias if this is differential between groups.

Participant follow-up methods varied by study. Studies in the EHNBPCCG generally followed-up participants via linkage to cancer registries until study closure. The national hospital admissions analysis in Chapter 7 "followed-up" men via national hospital admission data and mortality records. A hospital admission included any visit to a hospital which required a hospital bed, but excluded accident and emergency (A & E) attendance or outpatient bookings. Furthermore, this analysis was not linked to cancer registries. Therefore, the men identified in this analysis are likely to represent more severe forms of both prostate cancer and hormone-related diseases. UK Biobank followed-up patients via record linkage to a

combination of national death and cancer registries and hospital admission databases and therefore, will identify a wider range of prostate cancer diagnoses [283].

Record linkage may reduce the probability of loss to follow-up, but makes the estimation of any potential loss to follow-up difficult [559]. Data will not have been collected on patients who have emigrated or received treatment outside the individual study's country unless follow-up data included a self-report validation (some EHNBPCCG studies). Administrative errors will lead to some misclassification of prostate cancer status, but this is likely to be small [273].

Some studies also provided data on prostate tumour characteristics such as stage and grade information. It is possible that there was variation in tumour characterisation diagnostic practices between studies and over time, which may introduce bias [51, 560].

Reverse causation

Reverse causation occurs when the outcome precedes and causes the exposure. In all analyses men were excluded if they were known to have a history of prostate cancer.

To investigate the possibility of reverse causation, heterogeneity in the associations with prostate cancer diagnosis by time from baseline to diagnosis was investigated. Although there was no evidence of reverse causation in the associations with prostate cancer risk observed in Chapters 6-8, a high proportion of diagnosed prostate tumours are latent and/or grow very slowly [493, 494]. Median follow-up time ranged from 6.8-7.7 years which may be insufficient to rule-out reverse causation.

A strength of the analysis in Chapter 7 is that prostate tumours are not known to cause these hormone-related diseases and cannot have caused the genetic abnormality Klinefelter's syndrome. Therefore, reverse causation is less likely to account for the association with prostate cancer risk in these men.

9.3.3 The role of confounding

Only a relatively small number of risk factors for prostate cancer have been identified. Matching of cases to controls within the EHNBPCCG provided a way to control for other possible within-study confounders such as assay batch, age, region. Where possible, analyses also adjusted for potential confounders, which were chosen *a priori* on the basis of data availability and previous research.

For the analysis of hormone-related diseases and prostate cancer using routinely collected data (Chapter 7), participant characterisation data were limited. However, adjustment for these factors within the EHNBPCCG and in UK Biobank did not materially affect the associations with prostate cancer risk, so any possible residual confounding from these factors is not likely to have a large effect on the relative risk estimates from the analyses of the routinely collected data. Despite the range of data and study designs used in this thesis, and efforts to adjust for confounders, it is not possible to completely rule-out the possibility of residual confounding.

9.3.4 Summary of limitations

The research in this thesis represents a substantial contribution to the literature, particularly the observed threshold association between low free testosterone and a lower overall prostate cancer risk. The large size of the datasets analysed maximised the statistical power to study associations which may not previously have been detectable in smaller studies. The diversity of the study designs enabled a more thorough investigation into whether observed associations are robust across different contexts. However, the analyses presented in this thesis have a number of limitations and further research is also required to assess whether the associations observed are causal.

Given the high prevalence of latent prostate tumours and the widespread use of PSA testing, detection bias is an important limitation in assessing risk factors for prostate cancer. Androgen receptor signalling also partially regulates PSA synthesis, therefore low free testosterone may reduce the risk of tumour detection rather than incidence. Furthermore, aberrant hormone concentrations may also reflect poorer

health and lower socioeconomic status which may affect PSA testing attendance. Focusing on high-grade, advanced stage tumours or prostate cancer mortality may help reduce this bias, but this limits statistical power, which necessitates further efforts to collaborate and pool data.

9.4 Priorities for future research

9.4.1 Hormones in healthy and malignant prostate tissue

This thesis assumes that bound testosterone is unable to pass out of the bloodstream into the prostate tissue (i.e. the free hormone hypothesis [157]). Therefore, low free testosterone, rather than total, will reflect low intra-prostatic androgen activity. It is also assumed that higher circulating IGF-I concentrations will reflect higher intra-prostatic IGF-I concentrations [209, 210].

However, it is unclear to what extent peripheral free testosterone or IGF ("free" and/or bound to IGFBPs) concentrations are suitable surrogate markers of intra-prostatic hormone concentrations. It is particularly unclear to what extent circulating IGFBP concentrations may affect intra-prostatic IGF-I concentrations [167] (Chapter 2). Therefore, future research should validate the relationships between circulating free testosterone, IGF and IGFBP concentrations and intra-prostatic hormone concentrations using prostate tissue samples.

9.4.2 Triangulation

Triangulation describes the method of integrating results from several different approaches [561]. These different approaches include different study designs with differing sources of bias, as well as differing durations and timings of exposures. If the results from the different studies are consistent then this provides evidence for a robust association, which is more likely to be causal [561].

Analyses in this thesis tested the associations of free testosterone with prostate cancer risk in different contexts (Chapters 6 & 7), but more research is necessary. Further prospective analyses in the UK Biobank and Mendelian randomization analyses (outlined below) represent further opportunities to assess causality.

Prospective analyses in UK Biobank

The EHNBPCCG analysis provided the first large-scale prospective epidemiological evidence of an association between free testosterone and prostate cancer risk. Therefore, it is particularly important to replicate these findings in a large, standardised dataset. The UK Biobank recently released biomarker panel data (including circulating testosterone, SHBG, albumin and IGF-I concentrations), which were measured in samples collected at baseline across the entire cohort using standardised methods [282, 284]. The UK Biobank dataset also contains detailed participant characterisation data regarding medication usage, co-morbidities and health information and repeat measurements are available in a subset of approximately 9,000 men [291]. At present, there have been approximately 5,500 incident prostate cancer cases and 300 prostate cancer deaths in men with hormone measurement data.

This is the largest collection of hormones and prostate cancer status available in a single study, and UK Biobank therefore provides a unique opportunity to further explore the findings reported in Chapters 6 and 7. This will enable more robust estimates of relative risks and allow corroboration of the shape of the associations between circulating free testosterone concentrations and prostate cancer risk.

Mendelian randomization

In conventional observational prospective cohort studies, it is difficult to eliminate residual confounding. Germline genetic associations with hormone concentrations are not influenced by environmental factors and alleles are randomly allocated at conception (Figure 9.1). By investigating whether the genetic variants that are associated with endogenous hormone concentrations are also associated with prostate cancer risk, it may be possible to explore whether associations observed in Chapters 6 and 7 may be causal [562].

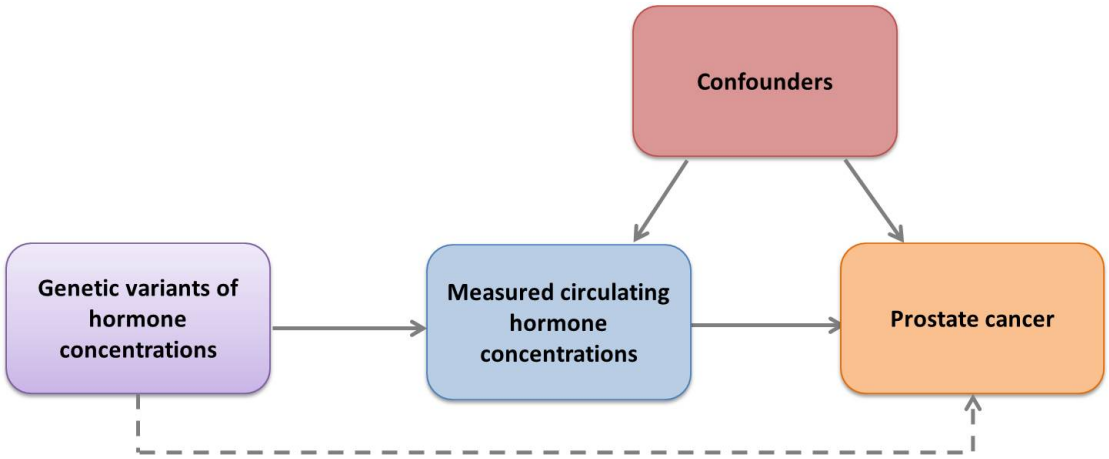


Figure 9.1: The Mendelian randomization approach to investigating free testosterone and prostate cancer risk

Although a Mendelian randomization analysis has been carried out to investigate the associations of genetically determined IGFs and IGFBPs with prostate cancer risk [219], potential pleiotropy of the genetic instruments meant it was not possible to identify the specific hormone/binding protein driving the association. Moreover, the association of free testosterone with prostate cancer risk has not been investigated using this method.

The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium contains genetic data from ∼45,000 prostate cancer cases (including ∼6,000 aggressive disease) from over 100 studies worldwide [563]. While Mendelian randomization is unaffected by environmental confounders and may reflect life-long genetic exposure, genetic instruments often predict <5% of the variance, and so these studies generally have a lower power than conventional observational studies [564]. Although it is possible to investigate non-linear associations using Mendelian randomization [565, 566], this may further limit statistical power. Therefore, still more data pooling and collaboration may be necessary to substantiate the threshold association between free testosterone and prostate cancer risk observed in Chapter 6.

9.4.3 Mediation analysis

If triangulation frameworks consistently implicate the roles of circulating free testosterone and IGF-I concentrations in prostate cancer risk, then mediation analysis may provide insights into possible prostate cancer preventative strategies. Analyses in Chapters 4 and 5 indicated that a range of anthropometric, sociodemographic and behavioural factors are associated with circulating free testosterone and IGF-I concentrations, indicating that concentrations of these hormones may be modifiable. Therefore, some of these observed associations may play a role in conferring prostate cancer risk, but this was not directly investigated in this thesis.

Mediation analysis in large standardised datasets with well-measured participant characterisation data such as UK Biobank will further substantiate whether the factors investigated in this thesis and other lifestyle factors (such as diet and physical activity) may play a role in prostate cancer risk via their associations with circulating hormone concentrations.

9.4.4 Aggressive tumour subtypes

Men diagnosed with high-grade prostate tumours have a poor prognosis [567]. However, a high proportion of diagnosed prostate tumours are localised and low-grade, and may never require clinical treatment [20, 22, 33, 36]. Many of these low-grade tumours are diagnosed due a PSA test, rather than the onset on clinical symptoms. More clinically aggressive tumour subtypes may be less susceptible to this detection bias, and the prevention of more clinically relevant tumours will have a greater public health impact. Therefore, differentiating between tumour subtypes is important in investigating potential prostate cancer risk factors.

More aggressive prostate tumour subtypes are less common, therefore large studies with longer durations of follow-up and with tumour data are required. Analysis in UK Biobank will provide greater power to examine further associations by tumour characteristics, with tumour stage and grade data which are due to be released at the end of 2019, and additional prostate cancer cases and deaths accruing as the cohort ages.

9.4.5 Collaboration and data pooling

As described above, individual prospective cohort studies may be underpowered to detect associations, particularly with clinically aggressive tumour subtypes and if the associations are non-linear. Therefore, the continuation of data pooling efforts including the EHNBPCCG and the PRACTICAL consortium are fundamental to investigating risk factors for clinically significant prostate cancer.

9.4.6 Implications for prostate cancer prevention

Reviews of the epidemiology and aetiology of prostate cancer have concluded that there are no well-established modifiable risk factors [90].

The studies described in this thesis add to the evidence that prostate cancer risk may be lower in men with low circulating free testosterone concentrations, although there was some evidence to support an increased risk of aggressive disease. This possible elevated risk of this more clinically relevant disease warrants further investigation before it is possible to infer prostate cancer preventative strategies, particularly because clinical trials which aimed to lower prostate cancer risk via reductions in intra-prostatic androgen concentrations reported similar associations to those observed in this thesis [8–10].

Results from this thesis also support the hypothesis that higher IGF-I concentrations may increase prostate cancer risk. Analysis in Chapter 5 also identified potential lifestyle and anthropometric determinants of the IGF system which might be amenable to lifestyle interventions in order to reduce risk. The strongest modifiable correlate of this system was obesity, which was associated with substantially lower IGFBP-1 and IGFBP-2 concentrations, which also have an inverse relationship with prostate cancer risk [14], possibly as these binding proteins limit IGF-I signalling. However, further research should be conducted to investigate whether the large impact of obesity on IGFBP-1 and IGFBP-2 concentrations is relevant for the association of obesity with aggressive prostate cancer risk [92].

Further studies should also examine other potential modifiable determinants of these hormones, including the role of physical activity and dietary factors in order to inform other possible prostate cancer prevention strategies.

9.5 Conclusion

The findings from this thesis support the roles of endogenous testosterone and IGF-I in prostate cancer development. The use of the pooled EHNBPCCG dataset was fundamental to examining complex non-linear associations, which previous single epidemiological studies may have been underpowered to detect. This thesis also identified other potential anthropometric, sociodemographic and behavioural factors that may affect prostate cancer risk via alterations in hormone concentrations.

The observed prospective association between low free testosterone and prostate cancer risk was novel, therefore more detailed analyses are required in large standardised cohort studies to corroborate the shape of the association and to investigate the possible increased risk of high-grade prostate cancer. Moreover, Mendelian randomization will be a useful further step towards assessing whether the observed associations may be causal.

Appendices

A
Appendix 1

A.1 Alternative free hormone calculations

A.1.1 Free testosterone

The Sartorius empirical formula [306]:

Free testosterone (pmol/L)

$$[FE] = 24.00314 \times T/log_{10}S - 0.04599 \times T^2 \quad \text{(A.1)}$$

Where:

[T] = total testosterone concentration (mol/L)

[S] = SHBG concentration (mol/L)

A.1.2 Free estradiol

Alternative methods for estimating free estradiol are available which also incorporate free testosterone concentrations into the formula, and may therefore be a more accurate measure of free estradiol [163, 346]:

Free estradiol (pmol/L)

$$[FE] = 1.0e^{12} \times (\sqrt{((N/k_{\mathrm{E}} - [E] + [S])^2/(4N^2) + k_{\mathrm{E}}/(Nk_{\mathrm{E}}))} - (N/k_{\mathrm{E}} - [E] + [S])/(2N)) \tag{A.2}$$

Where:

N = 1 + [A]k_{A}

[E] = total estradiol concentration (mol/L)

[T] = total testosterone concentration (mol/L)

[FT] = free testosterone concentration (mol/L)

[BT] = T - FT

[S] = SHBG concentration (mol/L) - [BT]

[A] = albumin concentration (assuming 6.2 x 10^{-4} L/mol)

k_{E} = association constant of estradiol to SHBG (assuming 0.6 x 10^{9})

k_{A}= association constant of estradiol to albumin (assuming 0.61 x 10^{4})

References

[1] F. Bray et al. “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA Cancer J Clin* 68.6 (2018), pp. 394–424.

[2] J. Ferlay et al. *Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer*. URL: `https://gco.iarc.fr/today`.

[3] F.C. Hamdy and I. Eardley. *Oxford textbook of urological surgery*. Textbook of urological surgery. Oxford: Oxford University Press, 2017.

[4] M. Pollard, P. H. Luckert, and M. A. Schmidt. “Induction of prostate adenocarcinomas in Lobund Wistar rats by testosterone”. In: *Prostate* 3.6 (1982), pp. 563–8.

[5] M. C. Bosland. “Testosterone treatment is a potent tumor promoter for the rat prostate”. In: *Endocrinology* 155.12 (2014), pp. 4629–33.

[6] H. Bu et al. “Putative prostate cancer risk SNP in an androgen receptor-binding site of the melanophilin gene illustrates enrichment of risk SNPs in androgen receptor target sites”. In: *Hum Mutat* 37.1 (2016), pp. 52–64.

[7] G. Thomas et al. “Multiple loci identified in a genome-wide association study of prostate cancer”. In: *Nat Genet* 40.3 (2008), pp. 310–315.

[8] I. Thompson et al. “Long-term survival of participants in the Prostate Cancer Prevention Trial”. In: *N Engl J Med* 369.7 (2013), pp. 603–610.

[9] G. L. Andriole et al. “Effect of dutasteride on the risk of prostate cancer”. In: *N Engl J Med* 362.13 (2010), pp. 1192–1202.

[10] I. M. Thompson et al. “The influence of finasteride on the development of prostate cancer”. In: *N Engl J Med* 349.3 (2003), pp. 215–24.

[11] N. Mottet et al. “EAU-ESTRO-SIOG Guidelines on prostate cancer. Part 1: Screening, diagnosis, and local treatment with curative intent”. In: *Eur Urol* 71.4 (2017), pp. 618–629.

[12] C. Huggins, Jr. Stevens R. E., and C. V. Hodges. “Studies on prostate cancer: II. The effects of castration on advanced carcinoma of the prostate gland”. In: *Archives of Surgery* 43.2 (1941), pp. 209–223.

[13] A. W. Roddam et al. “Endogenous sex hormones and prostate cancer: a collaborative analysis of 18 prospective studies”. In: *J Natl Cancer Inst* 100.3 (2008), pp. 170–83.

[14] R. C. Travis et al. “A meta-analysis of individual participant data reveals an association between circulating levels of IGF-I and prostate cancer risk”. In: *Cancer Res* 76.8 (2016), pp. 2288–300.

[15] T. Lopez and D. Hanahan. “Elevated levels of IGF-1 receptor convey invasive and metastatic capability in a mouse model of pancreatic islet tumorigenesis”. In: *Cancer Cell* 1.4 (2002), pp. 339–53.

[16] C. S. Mantzoros et al. “Insulin-like growth factor 1 in relation to prostate cancer and benign prostatic hyperplasia”. In: *Br J Cancer* 76.9 (1997), pp. 1115–8.

[17] P. Cohen et al. “Elevated levels of insulin-like growth factor-binding protein-2 in the serum of prostate cancer patients”. In: *J Clin Endocrinol Metab* 76.4 (1993), pp. 1031–5.

[18] T. M. Pham et al. “Relationship between serum levels of insulin-like growth factors and subsequent risk of cancer mortality: findings from a nested case-control study within the Japan Collaborative Cohort Study”. In: *Cancer Epidemiol* 34.3 (2010), pp. 279–84.

[19] A. Morgentaler and A. M. Traish. “Shifting the paradigm of testosterone and prostate cancer: the saturation model and the limits of androgen-dependent growth”. In: *Eur Urol* 55.2 (2009), pp. 310–20.

[20] S. Loeb et al. “Overdiagnosis and overtreatment of prostate cancer”. In: *Euro Urol* 65.6 (2014), pp. 1046–1055.

[21] W. J. Catalona et al. “Measurement of prostate-specific antigen in serum as a screening test for prostate cancer”. In: *N Engl J Med* 324.17 (1991), pp. 1156–1161.

[22] H. Carter. “Prostate-specific antigen (PSA) screening for prostate cancer: Revisiting the evidence”. In: *JAMA* 319.18 (2018), pp. 1866–1868.

[23] A. L. Potosky et al. “The role of increasing detection in the rising incidence of prostate cancer”. In: *JAMA* 273.7 (1995), pp. 548–552.

[24] K. A. Houston et al. “Trends in prostate cancer incidence rates and prevalence of prostate specific antigen screening by socioeconomic status and regions in the United States, 2004 to 2013”. In: *J Urol* 199.3 (2018), pp. 676–682.

[25] M. M. Center et al. “International variation in prostate cancer incidence and mortality rates”. In: *Eur Urol* 61.6 (2012), pp. 1079–1092.

[26] M. C. Wong et al. “Global incidence and mortality for prostate cancer: Analysis of temporal patterns and trends in 36 countries”. In: *Eur Urol* 70.5 (2016), pp. 862–874.

[27] D. Hashim et al. “The global decrease in cancer mortality: trends and disparities”. In: *Ann Oncol* 27.5 (2016), pp. 926–933.

[28] S. L. Chen et al. “Prostate cancer mortality-to-incidence ratios are associated with cancer care disparities in 35 countries”. In: *Sci Rep* 7 (2017), p. 40003.

[29] I. J. Powell et al. "Evidence supports a faster growth rate and/or earlier transformation to clinically significant prostate cancer in black than in white American men, and influences racial progression and mortality disparity". In: *J Urol* 183.5 (2010), pp. 1792–6.

[30] SEER. *Surveillance, Epidemiology, and End Results (SEER) Program 1973-2014*. URL: `www.seer.cancer.gov`.

[31] T. A. Stamey et al. "Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate". In: *N Engl J Med* 317.15 (1987), pp. 909–16.

[32] A. M. Wolf et al. "American Cancer Society guideline for the early detection of prostate cancer: update 2010". In: *CA Cancer J Clin* 60.2 (2010), pp. 70–98.

[33] E. A. Platz, A. M. De Marzo, and E. Giovannucci. "Prostate cancer association studies: pitfalls and solutions to cancer misclassification in the PSA era". In: *J Cell Biochem* 91.3 (2004), pp. 553–71.

[34] F. H. Schröder et al. "Screening and prostate-cancer mortality in a randomized European study". In: *N Engl J Med* 360.13 (2009), pp. 1320–1328.

[35] G.L. Andriole et al. "Mortality results from a randomized prostate-cancer screening trial". In: *N Engl J Med* 360.13 (2009), pp. 1310–1319.

[36] P. F. Pinsky et al. "Extended follow-up for prostate cancer incidence and mortality among participants in the Prostate, Lung, Colorectal and Ovarian randomized cancer screening trial". In: *BJU Int* 123.5 (2019), pp. 854–860.

[37] R. M. Martin et al. "Effect of a low-intensity psa-based screening intervention on prostate cancer mortality: The cap randomized clinical trial". In: *JAMA* 319.9 (2018), pp. 883–895.

[38] A. Tsodikov et al. "Reconciling the Effects of Screening on Prostate Cancer Mortality in the ERSPC and PLCO Trials". In: *Ann Intern Med* 167.7 (2017), pp. 449–455.

[39] K. A. O. Tikkinen et al. "Prostate cancer screening with prostate-specific antigen (PSA) test: a clinical practice guideline". In: *BMJ* 362 (2018), k3581.

[40] M. W. Drazer, D. Huo, and S. E. Eggener. "National prostate cancer screening rates after the 2012 US Preventive Services Task Force recommendation discouraging prostate-specific antigen–based screening". In: *J Clin Oncol* 33.22 (2015), pp. 2416–2423.

[41] J. Melia, S. Moss, and L. Johns. "Rates of prostate-specific antigen testing in general practice in England and Wales in asymptomatic and symptomatic patients: a cross-sectional study". In: *BJU Int* 94.1 (2004), pp. 51–6.

[42] N. Williams et al. "Prostate-specific antigen testing rates remain low in UK general practice: a cross-sectional study in six English cities". In: *BJU Int* 108.9 (2011), pp. 1402–8.

[43] G. J. Young et al. "Prostate-specific antigen (PSA) testing of men in UK general practice: a 10-year longitudinal cohort study". In: *BMJ Open* 7.10 (2017), e017729.

[44] T. J. Littlejohns et al. “Lifestyle factors and prostate-specific antigen (PSA) testing in UK Biobank: Implications for epidemiological research”. In: *Cancer Epidemiol* 45 (2016), pp. 40–46.

[45] F. H. Schröder et al. “The TNM classification of prostate cancer”. In: *Prostate* 21.S4 (1992), pp. 129–138.

[46] D.F. Gleason. “Histological grading and clinical staging of prostatic carcinoma”. In: *Urologic Pathology: The Prostate* 171 (1977).

[47] D. F. Gleason. “Histologic grading of prostate cancer: A perspective”. In: *Human Pathology* 23.3 (1992), pp. 273–279.

[48] N. Chen and Q. Zhou. “The evolving Gleason grading system”. In: *Chin J Cancer Res* 28.1 (2016), pp. 58–64.

[49] S. B. Edge and C. C. Compton. “The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM”. In: *Ann Surg Oncol* 17.6 (2010), pp. 1471–1474.

[50] K. Eichler et al. “Diagnostic value of systematic biopsy methods in the investigation of prostate cancer: a systematic review”. In: *J Urol* 175.5 (2006), pp. 1605–12.

[51] P. C. Albertsen et al. “Prostate Cancer and the Will Rogers Phenomenon”. In: *J Natl Cancer Inst* 97.17 (2005), pp. 1248–1253.

[52] J. I. Epstein et al. “The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System”. In: *Am J Surg Pathol* 40.2 (2016), pp. 244–52.

[53] C. A. Reichard and E. A. Klein. “Clinical and molecular rationale to retain the cancer descriptor for Gleason score 6 disease”. In: *Nat Rev Urol* 14.1 (2017), pp. 59–64.

[54] H. B. Carter, S. Piantadosi, and J. T. Isaacs. “Clinical evidence for and implications of the multistep development of prostate cancer”. In: *J Urol* 143.4 (1990), pp. 742–6.

[55] A. Pettersson et al. “Age at diagnosis and prostate cancer treatment and prognosis: a population-based cohort study”. In: *Ann Oncol* 29.2 (2018), pp. 377–385.

[56] M. M. Shen and C. Abate-Shen. “Molecular genetics of prostate cancer: new prospects for old challenges”. In: *Genes Dev* 24.18 (2010), pp. 1967–2000.

[57] M. K. Graham and A. Meeker. “Telomeres and telomerase in prostate cancer development and therapy”. In: *Nat Rev Urol* 14.10 (2017), pp. 607–619.

[58] A. Zinger, W. C. Cho, and A. Ben-Yehuda. “Cancer and aging - the inflammatory connection”. In: *Aging Dis* 8.5 (2017), pp. 611–627.

[59] A. B. Araujo and G. A. Wittert. “Endocrinology of the aging male”. In: *Best Pract Res Clin Endocrinol Metab* 25.2 (2011), pp. 303–319.

[60] T. Finkel, M. Serrano, and M. A. Blasco. "The common biology of cancer and ageing". In: *Nature* 448.7155 (2007), pp. 767–74.

[61] M. Kiciński, J. Vangronsveld, and T. S. Nawrot. "An epidemiological reappraisal of the familial aggregation of prostate cancer: A meta-analysis". In: *PLOS ONE* 6.10 (2011), e27130.

[62] L. A. Mucci et al. "Familial risk and heritability of cancer among twins in Nordic countries". In: *JAMA* 315.1 (2016), pp. 68–76.

[63] B. Meiser et al. "Prostate cancer screening in men with a family history of prostate cancer: the role of partners in influencing men's screening uptake". In: *Urology* 70.4 (2007), pp. 738–42.

[64] S. Benafif, Z. Kote-Jarai, and R. A. Eeles. "A review of prostate cancer genome-wide association studies (GWAS)". In: *Cancer Epidemiol Biomarkers Prev* 27.8 (2018), pp. 845–857.

[65] F. R. Schumacher et al. "Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci". In: *Nature Genetics* 50.7 (2018), pp. 928–936.

[66] Z. Shang et al. "Germline homeobox B13 (HOXB13) G84E mutation and prostate cancer risk in European descendants: a meta-analysis of 24,213 cases and 73, 631 controls". In: *Eur Urol* 64.1 (2013), pp. 173–6.

[67] J. D. Norris et al. "The homeodomain protein HOXB13 regulates the cellular response to androgens". In: *Mol Cell* 36.3 (2009), pp. 405–16.

[68] H. Chen et al. "Systematic enrichment analysis of potentially functional regions for 103 prostate cancer risk-associated loci". In: *Prostate* 75.12 (2015), pp. 1264–76.

[69] Y. Lu et al. "Functional annotation of risk loci identified through genome-wide association studies for prostate cancer". In: *Prostate* 71.9 (2011), pp. 955–963.

[70] W. B. Isaacs and J. Xu. "Current progress and questions in germline genetics of prostate cancer". In: *Asian J Urol* 6.1 (2019), pp. 3–9.

[71] R. L. Siegel, K. D. Miller, and A. Jemal. "Cancer Statistics, 2017". In: *CA Cancer J Clin* 67.1 (2017), pp. 7–30.

[72] C. E. DeSantis et al. "Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities". In: *CA Cancer J Clin* 66.4 (2016), pp. 290–308.

[73] Y. Ben-Shlomo et al. "The risk of prostate cancer amongst black men in the United Kingdom: The PROCESS cohort study". In: *Euro Urol* 53.1 (2008), pp. 99–105.

[74] A. Perez-Cornago et al. "Prospective investigation of risk factors for prostate cancer in the UK Biobank cohort study". In: *Br J Cancer* 117.10 (2017), pp. 1562–1571.

[75] D. Y. Lichtensztajn et al. "Prostate cancer risk profiles of Asian-American men: disentangling the effects of immigration status and race/ethnicity". In: *J Urol* 191.4 (2014), pp. 952–956.

[76] B. A. Mahal et al. "Prostate Cancer-Specific Mortality Across Gleason Scores in Black vs Nonblack Men". In: *Jama* 320.23 (2018), pp. 2479–2481.

[77] B. A. Miller et al. "Cancer incidence and mortality patterns among specific Asian and Pacific Islander populations in the U.S". In: *Cancer Causes Control* 19.3 (2008), pp. 227–256.

[78] R. W. Haile et al. "A review of cancer in U.S. Hispanic populations". In: *Cancer Prev Res (Phila)* 5.2 (2012), pp. 150–163.

[79] E. A. Miller, P. F. Pinsky, and D. Pierre-Victor. "The relationship between diabetes, prostate-specific antigen screening tests, and prostate cancer". In: *Cancer Causes Control* 29.10 (2018), pp. 907–914.

[80] G. F. Chao et al. "Asian Americans and prostate cancer: A nationwide population-based analysis". In: *Urol Oncol* 34.5 (2016), 233.e7–15.

[81] C. M. Zeigler-Johnson et al. "Genetic susceptibility to prostate cancer in men of African descent: implications for global disparities in incidence and outcomes". In: *Can J Urol* 15.1 (2008), pp. 3872–3882.

[82] M. L. Freedman et al. "Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men". In: *Proc Natl Acad Sci USA* 103.38 (2006), pp. 14068–73.

[83] T. R. Rebbeck. "Prostate cancer genetics: variation by race, ethnicity, and geography". In: *Semin Radiat Oncol* 27.1 (2017), pp. 3–10.

[84] D. Hatcher et al. "Molecular mechanisms involving prostate cancer racial disparity". In: *Am J Transl Res* 1.3 (2009), pp. 235–248.

[85] S. Rohrmann et al. "Serum estrogen, but not testosterone, levels differ between black and white men in a nationally representative sample of Americans". In: *J Clin Endocrinol Metab* 92.7 (2007), pp. 2519–25.

[86] R. K. Ross et al. "5-alpha-reductase activity and risk of prostate cancer among Japanese and US white and black males". In: *Lancet* 339.8798 (1992), pp. 887–9.

[87] A. S. Whittemore et al. "Prostate cancer in relation to diet, physical activity, and body size in blacks, whites, and Asians in the United States and Canada". In: *J Natl Cancer Inst* 87.9 (1995), pp. 652–61.

[88] L. S. Cook et al. "Incidence of adenocarcinoma of the prostate in Asian immigrants to the United States and their descendants". In: *J Urol* 161.1 (1999), pp. 152–5.

[89] D. M. Parkin et al. "Cancer in Africa 2012". In: *Cancer Epidemiol Biomarkers Prev* 23.6 (2014), pp. 953–66.

[90] WCRF/AICR. *World Cancer Research Fund/American Institute for Cancer Research continuous update project: diet, nutrition, physical activity, and prostate cancer*. URL: https://www.wcrf.org/sites/default/files/Prostate-Cancer-2014-Report.pdf.

[91] I. dos-Santos-Silva and R. Denholm. "Height and risk of adult cancers: A review". In: *Cur Epidemiol Rep* 3.3 (2016), pp. 191–200.

[92] A. Perez-Cornago et al. "Tall height and obesity are associated with an increased risk of aggressive prostate cancer: results from the EPIC cohort study". In: *BMC Med* 15.1 (2017), p. 115.

[93] D. M. Holman et al. "The association between adverse childhood experiences and risk of cancer in adulthood: A systematic review of the literature". In: *Pediatrics* 138.Suppl 1 (2016), S81–S91.

[94] S. Yakar et al. "Circulating levels of IGF-1 directly regulate bone growth and density". In: *J Clin Invest* 110.6 (2002), pp. 771–781.

[95] D. Albanes and M. Winick. "Are cell number and cell proliferation risk factors for cancer?" In: *J Natl Cancer Inst* 80.10 (1988), pp. 772–4.

[96] N. K. Khankari et al. "Association between adult height and risk of colorectal, lung, and prostate cancer: Results from meta-analyses of prospective studies and Mendelian randomization analyses". In: *PLOS Medicine* 13.9 (2016), e1002118.

[97] N. M. Davies et al. "The effects of height and BMI on prostate cancer incidence and mortality: a Mendelian randomization study in 20,848 cases and 20,214 controls from the PRACTICAL consortium". In: *Cancer Causes Control* 26.11 (2015), pp. 1603–16.

[98] S. Sutcliffe and G. A. Colditz. "Prostate cancer: is it time to expand the research focus to early-life exposures?" In: *Nat Rev Cancer* 13.3 (2013), pp. 208–518.

[99] C. Gao et al. "Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer". In: *Int J Epidemiol* 45.3 (2016), pp. 896–908.

[100] S. J. Freedland and E. A. Platz. "Obesity and prostate cancer: making sense out of apparently conflicting data". In: *Epidemiol Rev* 29 (2007), pp. 88–97.

[101] J. H. Fowke et al. "Association between biomarkers of obesity and risk of high-grade prostatic intraepithelial neoplasia and prostate cancer–evidence of effect modification by prostate size". In: *Cancer Lett* 328.2 (2013), pp. 345–52.

[102] E. Giovannucci. "A framework to understand diet, physical activity, body weight, and cancer risk". In: *Cancer Causes Control* 29.1 (2018), pp. 1–6.

[103] M. Poddar, Y. Chetty, and V. T. Chetty. "How does obesity affect the endocrine system? A narrative review". In: *Clin Obes* 7.3 (2017), pp. 136–144.

[104] E. E. Calle and R. Kaaks. "Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms". In: *Nat Rev Cancer* 4 (2004), p. 579.

[105] K. T. McVary. "BPH: epidemiology and comorbidities". In: *Am J Manag Care* 12.5 Suppl (2006), S122–8.

[106] D. G. Bostwick et al. "The association of benign prostatic hyperplasia and cancer of the prostate". In: *Cancer* 70.1 Suppl (1992), pp. 291–301.

[107] X. Dai et al. "Benign prostatic hyperplasia and the risk of prostate cancer and bladder cancer: A meta-analysis of observational studies". In: *Medicine* 95.18 (2016), e3493.

[108] J. M. Schenk et al. "Association of symptomatic benign prostatic hyperplasia and prostate cancer: Results from the Prostate Cancer Prevention Trial". In: *Am J Epidemiol* 173.12 (2011), pp. 1419–1428.

[109] D. Bansal et al. "Type 2 diabetes and risk of prostate cancer: a meta-analysis of observational studies". In: *Prostate Cancer Prostatic Dis* 16.2 (2013), 151–8, s1.

[110] K. Fall et al. "Diabetes mellitus and prostate cancer risk; a nationwide case-control study within PCBaSe Sweden". In: *Cancer Epidemiol Biomarkers Prev* 22.6 (2013), pp. 1102–9.

[111] K. K. Tsilidis et al. "Diabetes mellitus and risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition". In: *Int J Cancer* 136.2 (2015), pp. 372–81.

[112] R. Dankner et al. "Diabetes, prostate cancer screening and risk of low- and high-grade prostate cancer: an 11 year historical population follow-up study of more than 1 million men". In: *Diabetologia* 59.8 (2016), pp. 1683–91.

[113] D. M. Werny, M. Saraiya, and E. W. Gregg. "Prostate-specific antigen values in diabetic and nondiabetic US men, 2001-2002". In: *Am J Epidemiol* 164.10 (2006), pp. 978–83.

[114] N. Friedrich et al. "The association between IGF-I and insulin resistance: a general population study in Danish adults". In: *Diabetes Care* 35.4 (2012), pp. 768–773.

[115] B. A. Calton et al. "History of diabetes mellitus and subsequent prostate cancer risk in the NIH-AARP Diet and Health Study". In: *Cancer Causes Control* 18.5 (2007), pp. 493–503.

[116] C. Wu et al. "Diabetes and prostate cancer risk in the REDUCE trial". In: *Prostate Cancer Prostatic Dis* 14.4 (2011), pp. 326–31.

[117] S. M. Crusz and F. R. Balkwill. "Inflammation and cancer: advances and new agents". In: *Nat Rev Clin Oncol* 12.10 (2015), pp. 584–96.

[118] H. Gonzalez, C. Hagerling, and Z. Werb. "Roles of the immune system in cancer: from tumor initiation to metastatic progression". In: *Genes Dev* 32.19-20 (2018), pp. 1267–1284.

[119] E. A. Platz and A. M. De Marzo. "Epidemiology of inflammation and prostate cancer". In: *J Urol* 171.2, Supplement (2004), S36–S40.

[120] R. Arthur et al. "Serum inflammatory markers in relation to prostate cancer severity and death in the Swedish AMORIS study". In: *Int J Cancer* 142.11 (2018), pp. 2254–2262.

[121] K. Beckmann et al. "Chronic inflammatory diseases, anti-inflammatory medications and risk of prostate cancer: a population-based case-control study". In: *BMC Cancer* 19.1 (2019), p. 612.

[122] A. Vasconcelos et al. "Dairy products: Is there an impact on promotion of prostate cancer? A review of the literature". In: *Front Nutr* 6 (2019), pp. 62–62.

[123] R. P. Heaney et al. "Dietary changes favorably affect bone remodeling in older adults". In: *J Am Diet Assoc* 99.10 (1999), pp. 1228–33.

[124] M. D. Holmes et al. "Dietary correlates of plasma insulin-like growth factor I and insulin-like growth factor binding protein 3 concentrations". In: *Cancer Epidemiol Biomarkers Prev* 11.9 (2002), pp. 852–61.

[125] M. Tseng et al. "Dairy, calcium, and vitamin D intakes and prostate cancer risk in the National Health and Nutrition Examination Epidemiologic Follow-up Study cohort". In: *Am J Clin Nutr* 81.5 (2005), pp. 1147–54.

[126] E. Giovannucci. "Dietary influences of 1,25(OH)2 vitamin D in relation to prostate cancer: a hypothesis". In: *Cancer Causes Control* 9.6 (1998), pp. 567–82.

[127] G. G. Schwartz et al. "1,25-Dihydroxy-16-ene-23-yne-vitamin D3 and prostate cancer cell proliferation in vivo". In: *Urology* 46.3 (1995), pp. 365–9.

[128] J. Yarmolinsky et al. "Circulating selenium and prostate cancer risk: a Mendelian randomization analysis". In: *J Natl Cancer Inst* 110.9 (2018), pp. 1035–1038.

[129] M. N. I. Barclay, A. MacPherson, and J. Dixon. "Selenium content of a range of UK foods". In: *J Food Comp Analysis* 8.4 (1995), pp. 307–318.

[130] N. E. Allen et al. "Selenium and Prostate Cancer: Analysis of Individual Participant Data From Fifteen Prospective Studies". In: *J Natl Cancer Inst* 108.11 (2016).

[131] S. M. Lippman et al. "Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the Selenium and Vitamin E Cancer Prevention Trial (SELECT)". In: *JAMA* 301.1 (2009), pp. 39–51.

[132] R. M. Winkels et al. "Bioavailability of food folates is 80 percent of that of folic acid". In: *Am J Clin Nutr* 85.2 (2007), pp. 465–73.

[133] K. J. Rycyna, D. J. Bacich, and D. S. O'Keefe. "Opposing roles of folate in prostate cancer". In: *Urology* 82.6 (2013), pp. 1197–203.

[134] M. Fenech. "Folate (vitamin B9) and vitamin B12 and their function in the maintenance of nuclear and mitochondrial genome integrity". In: *Mutat Res* 733.1-2 (2012), pp. 21–33.

[135] P. J. Stover and M. S. Field. "Trafficking of intracellular folates". In: *Adv Nutr* 2.4 (2011), pp. 325–331.

[136] C. L. Krumdieck and P. N. Howard-Peebles. "On the nature of folic-acid-sensitive fragile sites in human chromosomes: an hypothesis". In: *Am J Med Genet* 16.1 (1983), pp. 23–8.

[137] A. J. Price et al. "Circulating folate and vitamin B(12) and risk of prostate cancer: A collaborative analysis of individual participant data from six cohorts including 6875 cases and 8104 controls". In: *Euro Urol* 70.6 (2016), pp. 941–951.

[138] S. E. Vollset et al. "Effects of folic acid supplementation on overall and site-specific cancer incidence during the randomised trials: meta-analyses of data on 50,000 individuals". In: *Lancet* 381.9871 (2013), pp. 1029–36.

[139] S. de Vogel et al. "Serum folate and vitamin B12 concentrations in relation to prostate cancer risk–a Norwegian population-based nested case-control study of 3000 cases and 3000 controls within the JANUS cohort". In: *Int J Epidemiol* 42.1 (2013), pp. 201–10.

[140] S. M. Collin et al. "Association of folate-pathway gene polymorphisms with the risk of prostate cancer: a population-based nested case-control study, systematic review, and meta-analysis". In: *Cancer Epidemiol Biomarkers Prev* 18.9 (2009), pp. 2528–39.

[141] X. L. Li and J. H. Xu. "MTHFR polymorphism and the risk of prostate cancer: a meta-analysis of case–control studies". In: *Prostate Cancer Prostatic Dis* 15 (2012), p. 244.

[142] S. Swami, A. V. Krishnan, and D. Feldman. "Vitamin D metabolism and action in the prostate: implications for health and disease". In: *Mol Cell Endocrinol* 347.1-2 (2011), pp. 61–9.

[143] K. Kostner et al. "The relevance of vitamin D receptor (VDR) gene polymorphisms for cancer: a review of the literature". In: *Anticancer Res* 29.9 (2009), pp. 3511–36.

[144] M. F. Holick. "Vitamin D status: measurement, interpretation, and clinical application". In: *Ann Epidemiol* 19.2 (2009), pp. 73–78.

[145] R. C. Travis et al. "A collaborative analysis of individual participant data from 19 prospective studies assesses circulating vitamin D and prostate cancer risk". In: *Cancer Research* 79.1 (2019), pp. 274–285.

[146] X. Jiang et al. "Circulating vitamin D concentrations and risk of breast and prostate cancer: a Mendelian randomization study". In: *Int J Epidemiol* (2018).

[147] A.B. Hoffmann. *Sex hormones: development, regulation and disorders*. Endocrinology research and clinical developments. New York: Nova Science Publishers, 2011.

[148] R. A. Davey and M. Grossmann. “Androgen receptor structure, function and biology: From bench to bedside”. In: *Clin Biochem Rev* 37.1 (2016), pp. 3–15.

[149] E. E. Baulieu and P. Mauvais-Jarvis. “Studies on Testosterone Metabolism: I. Conversion of testosterone-17α-3h to 5α- and 5β-androstane-3α, 17β-diol-17α-3h: A new "17β-hydroxyl pathway"”. In: *J Bio Chem* 239.5 (1964), pp. 1569–1577.

[150] C. W. Duan and L. Xu. “Testosterone and androstanediol glucuronide among men in NHANES III”. In: *BMC Public Health* 18.1 (2018), p. 339.

[151] P. S. Cooke et al. “Estrogens in male physiology”. In: *Physiol Rev* 97.3 (2017), pp. 995–1043.

[152] R. S. Rosenfeld, B. J. Rosenberg, and L. Hellman. “Direct analysis of dehydroisoandrosterone in plasma”. In: *Steroids* 25.6 (1975), pp. 799–805.

[153] S. Legrain et al. “Dehydroepiandrosterone replacement administration: Pharmacokinetic and pharmacodynamic studies in healthy elderly subjects”. In: *J Clin Endocrinol Metab* 85.9 (2000), pp. 3208–3217.

[154] D. El-Maouche and A. Dobs. “Chapter 60 - Testosterone replacement therapy in men and women”. In: *Principles of Gender-Specific Medicine (Second Edition)*. Ed. by Marianne J. Legato. San Diego: Academic Press, 2010, pp. 737–760.

[155] D. I. Spratt et al. “Neuroendocrine-gonadal axis in men: frequent sampling of LH, FSH, and testosterone”. In: *Am J Physiol* 254.5 (1988), E658–E666.

[156] F. Labrie et al. “Intracrinology: role of the family of 17 beta-hydroxysteroid dehydrogenases in human physiology and disease”. In: *J Mol Endocrinol* 25.1 (2000), pp. 1–16.

[157] C. M. Mendel. “The free hormone hypothesis: a physiologically based mathematical model”. In: *Endocr Rev* 10.3 (1989), pp. 232–74.

[158] D. J. Handelsman. “Free testosterone: pumping up the tires or ending the free ride?” In: *Endocr Rev* 38.4 (2017), pp. 297–301.

[159] A. L. Goldman et al. “A reappraisal of testosterone’s binding in circulation: physiological and clinical implications”. In: *Endocr Rev* (2017).

[160] J. F. Dunn, B. C. Nisula, and D. Rodbard. “Transport of steroid hormones: binding of 21 endogenous steroids to both testosterone-binding globulin and corticosteroid-binding globulin in human plasma”. In: *J Clin Endocrinol Metab* 53.1 (1981), pp. 58–68.

[161] D. M. Selva and G. L. Hammond. “Thyroid hormones act indirectly to increase sex hormone-binding globulin production by liver via hepatocyte nuclear factor-4alpha”. In: *J Mol Endocrinol* 43.1 (2009), pp. 19–27.

[162] M. Pugeat et al. “Sex hormone-binding globulin gene expression in the liver: Drugs and the metabolic syndrome”. In: *Mol Cell Endocrinol* 316.1 (2010), pp. 53–59.

[163] R. Sodergard et al. “Calculation of free and bound fractions of testosterone and estradiol-17 beta to human plasma proteins at body temperature”. In: *J Steroid Biochem* 16.6 (1982), pp. 801–10.

[164] M. R. Laurent et al. “Sex hormone-binding globulin regulation of androgen bioactivity in vivo: validation of the free hormone hypothesis”. In: *Sci Rep* 6 (2016), pp. 35539–35539.

[165] M. P. Czub et al. “Testosterone meets albumin – the molecular mechanism of sex hormone transport by serum albumins”. In: *Chem Sci* 10.6 (2019), pp. 1607–1618.

[166] D. LeRoith and C. T. Roberts. “The insulin-like growth factor system and cancer”. In: *Cancer Lett* 195 (2003).

[167] J. B. Allard and C. Duan. “IGF-binding proteins: why do they exist and why are there so many?” In: *Front Endocrinol (Lausanne)* 9 (2018), pp. 117–117.

[168] A. J. Varewijck and J. A. Janssen. “Insulin and its analogues and their affinities for the IGF1 receptor”. In: *Endocr Relat Cancer* 19.5 (2012), F63–75.

[169] J. I. Jones and D. R. Clemmons. “Insulin-like growth factors and their binding proteins: biological actions”. In: *Endocr Rev* 16.1 (1995), pp. 3–34.

[170] P. De Meyts et al. “The insulin-like growth factor-I receptor”. In: *Horm Res* 42.4-5 (1994), pp. 152–169.

[171] S. Benyoucef et al. “Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation”. In: *Biochem J* 403.3 (2007), pp. 603–613.

[172] O.G.P. Isaksson et al. “Metabolic functions of liver-derived (endocrine) insulin-like growth factor I”. In: *Hormone Research in Paediatrics* 55.Suppl. 2 (2001), pp. 18–21.

[173] C. M. Perks and J. M. P. Holly. *Insulin-like growth factors in health and disease*. Endocrinology and metabolism clinics of North America. Philadelphia: Saunders, 2012.

[174] S. Rajaram, D. J. Baylink, and S. Mohan. “Insulin-like growth factor-binding proteins in serum and other biological fluids: regulation and functions”. In: *Endocr Rev* 18.6 (1997), pp. 801–31.

[175] S. M. Firth and R. C. Baxter. “Cellular actions of the insulin-like growth factor binding proteins”. In: *Endocr Rev* 23.6 (2002), pp. 824–54.

[176] J. P. Thissen, J. M. Ketelslegers, and L. E. Underwood. “Nutritional regulation of the insulin-like growth factors”. In: *Endocr Rev* 15.1 (1994), pp. 80–101.

[177] R. C. Bunn and J. L. Fowlkes. “Insulin-like growth factor binding protein proteolysis”. In: *Trends Endocrinol Metab* 14.4 (2003), pp. 176–181.

[178] E. Giovannucci. "Nutrition, insulin, insulin-like growth factors and cancer". In: *Horm Metab Res* 35.11-12 (2003), pp. 694–704.

[179] L.A. Bach. "40 Years of IGF1: IGF-binding proteins". In: *J Mol Endocrinol* 61.1 (2018), T11.

[180] J. E. McNeal et al. "Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread". In: *Am J Surg Pathol* 12.12 (1988), pp. 897–906.

[181] Y. Zhou, E. C. Bolton, and J. O. Jones. "Androgens and androgen receptor signaling in prostate tumorigenesis". In: *J Mol Endocrinol* 54.1 (2015), R15–29.

[182] A. S. Goldstein et al. "Identification of a cell of origin for human prostate cancer". In: *Science* 329.5991 (2010), pp. 568–71.

[183] C. T. Wu et al. "Increased prostate cell proliferation and loss of cell differentiation in mice lacking prostate epithelial androgen receptor". In: *Proc Natl Acad Sci USA* 104.31 (2007), pp. 12679–12684.

[184] A. M. De Marzo et al. "Stem cell features of benign and malignant prostate epithelial cells". In: *J Urol* 160.6 Pt 2 (1998), pp. 2381–92.

[185] H. Bonkhoff and K. Remberger. "Widespread distribution of nuclear androgen receptors in the basal cell layer of the normal and hyperplastic human prostate". In: *Virchows Arch A Pathol Anat Histopathol* 422.1 (1993), pp. 35–8.

[186] S. Yu et al. "Altered prostate epithelial development in mice lacking the androgen receptor in stromal fibroblasts". In: *Prostate* 72.4 (2012), pp. 437–49.

[187] H. F. English, N. Kyprianou, and J. T. Isaacs. "Relationship between DNA fragmentation and apoptosis in the programmed cell death in the rat prostate following castration". In: *Prostate* 15.3 (1989), pp. 233–50.

[188] Y. Sugimura, G. R. Cunha, and A. A. Donjacour. "Morphological and histological study of castration-induced degeneration and androgen-induced regeneration in the mouse prostate". In: *Biol Reprod* 34.5 (1986), pp. 973–83.

[189] K. B. Cleutjens et al. "Two androgen response regions cooperate in steroid hormone regulated activity of the prostate-specific antigen promoter". In: *J Biol Chem* 271.11 (1996), pp. 6379–88.

[190] L. S. Marks et al. "The interpretation of serum prostate specific antigen in men receiving 5alpha-reductase inhibitors: a review and clinical recommendations". In: *J Urol* 176.3 (2006), pp. 868–74.

[191] S. T. Page et al. "Persistent intraprostatic androgen concentrations after medical castration in healthy men". In: *J Clin Endocrinol Metab* 91.10 (2006), pp. 3850–6.

[192] S. T. Page et al. "Dihydrotestosterone administration does not increase intraprostatic androgen concentrations or alter prostate androgen action in healthy men: a randomized-controlled trial". In: *J Clin Endocrinol Metab* 96.2 (2011), pp. 430–7.

[193] E. A. Mostaghel et al. "Impact of male hormonal contraception on prostate androgens and androgen action in healthy men: a randomized, controlled trial". In: *J Clin Endocrinol Metab* 97.8 (2012), pp. 2809–17.

[194] J. P. Weber et al. "The influence of reversible androgen deprivation on serum prostate-specific antigen levels in men with benign prostatic hyperplasia". In: *J Urol* 141.4 (1989), pp. 987–92.

[195] J. C. Nickel. "Comparison of clinical trials with finasteride and dutasteride". In: *Rev Urol* 6.Suppl 9 (2004), S31–S39.

[196] R. S. Swerdloff et al. "Dihydrotestosterone: biochemistry, physiology, and clinical Implications of elevated blood levels". In: *Endocr Rev* 38.3 (2017), pp. 220–254.

[197] J. D. Wilson et al. "5α-Androstane-3α,17β-diol is formed in tammar wallaby pouch young testes by a pathway involving 5α-pregnane-3α,17α-diol-20-one as a key intermediate". In: *Endocrinology* 144.2 (2003), pp. 575–580.

[198] G. Shaw et al. "Role of the alternate pathway of dihydrotestosterone formation in virilization of the Wolffian ducts of the tammar wallaby, Macropus eugenii". In: *Endocrinology* 147.5 (2006), pp. 2368–73.

[199] T. M. Penning et al. "Human 3alpha-hydroxysteroid dehydrogenase isoforms (AKR1C1-AKR1C4) of the aldo-keto reductase superfamily: functional plasticity and tissue distribution reveals roles in the inactivation and formation of male and female sex hormones". In: *Biochem J* 351.Pt 1 (2000), pp. 67–77.

[200] D. R. Bauman et al. "Identification of the major oxidative 3α-hydroxysteroid dehydrogenase in human prostate that converts 5α-androstane-3α,17β-diol to 5α-dihydrotestosterone: A potential therapeutic target for androgen-dependent disease". In: *Mol Endocrinol* 20.2 (2006), pp. 444–458.

[201] M. V. Fiandalo et al. "Inhibition of dihydrotestosterone synthesis in prostate cancer by combined frontdoor and backdoor pathway blockade". In: *Oncotarget* 9.13 (2018), pp. 11227–11242.

[202] S. A. Shain and R. W. Boesel. "Saturation analysis of the binding of androgens, antiandrogens and estrogens by the cytoplasmic high affinity androgen receptor of the rat ventral prostate". In: *J Steroid Biochem* 6.1 (1975), pp. 43–50.

[203] B. S. Keenan et al. "Androgen receptor in human skin fibroblasts. Characterization of a specific 17beta-hydroxy-5alpha-androstan-3-one-protein complex in cell sonicates and nuclei". In: *Steroids* 25.4 (1975), pp. 535–52.

[204] M. Khera et al. "Changes in prostate specific antigen in hypogonadal men after 12 months of testosterone replacement therapy: support for the prostate saturation theory". In: *J Urol* 186.3 (2011), pp. 1005–11.

[205] P. Cohen et al. “Insulin-like growth factors (IGFs), IGF receptors, and IGF-binding proteins in primary cultures of prostate epithelial cells”. In: *J Clin Endocrinol Metab* 73.2 (1991), pp. 401–7.

[206] I. Hers, E. E. Vincent, and J. M. Tavare. “Akt signalling in health and disease”. In: *Cell Signal* 23.10 (2011), pp. 1515–27.

[207] K. Siddle. “Signalling by insulin and IGF receptors: supporting acts and new players”. In: *J Mol Endocrinol* 47.1 (2011), R1–10.

[208] C. Boudon et al. “Secretion of insulin-like growth factors and their binding proteins by human normal and hyperplastic prostatic cells in primary culture”. In: *J Clin Endocrinol Metab* 81.2 (1996), pp. 612–7.

[209] M. Pollak. “Insulin-like growth factor physiology and cancer risk”. In: *Eur J Cancer* 36.10 (2000), pp. 1224–1228.

[210] M. Pollak. “Insulin and insulin-like growth factor signalling in neoplasia”. In: *Nat Rev Cancer* 8 (2008).

[211] C. Huggins and C. V. Hodges. “Studies on prostatic cancer. I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate”. In: *Cancer Research* 1.4 (1941), pp. 293–297.

[212] S. Rinaldi et al. “Serum levels of IGF-I, IGFBP-3 and colorectal cancer risk: results from the EPIC cohort, plus a meta-analysis of prospective studies”. In: *Int J Cancer* 126.7 (2010), pp. 1702–15.

[213] T. J. Key et al. “Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies”. In: *Lancet Oncol* 11 (2010).

[214] J. A. Schmidt et al. “Insulin-like growth factor-I and risk of differentiated thyroid carcinoma in the European Prospective Investigation into Cancer and Nutrition”. In: *Cancer Epidemiol Biomarkers Prev* 23.6 (2014), pp. 976–985.

[215] G. Fürstenberger and H. J. Senn. “Insulin-like growth factors and cancer”. In: *Lancet Oncol* 3.5 (2002), pp. 298–302.

[216] M. E. Cox et al. “Insulin receptor expression by human prostate cancers”. In: *Prostate* 69.1 (2009), pp. 33–40.

[217] B. S. Schaffer et al. “Opposing roles for the insulin-like growth factor (IGF)-II and mannose 6-phosphate (Man-6-P) binding activities of the IGF-II/Man-6-P receptor in the growth of prostate cancer cells”. In: *Endocrinology* 144.3 (2003), pp. 955–966.

[218] T.U. Ahearn et al. “Expression of IGF/insulin receptor in prostate cancer tissue and progression to lethal disease”. In: *Carcinogenesis* 39.12 (2018), pp. 1431–1437.

[219] C. Bonilla et al. "Assessing the role of insulin-like growth factors and binding proteins in prostate cancer using Mendelian randomization: Genetic variants as instruments for circulating levels". In: *Int J Cancer* 139.7 (2016), pp. 1520–33.

[220] K. Zu et al. "Protein expression of PTEN, insulin-like growth factor I receptor (IGF-IR), and lethal prostate cancer: a prospective study". In: *Cancer Epidemiol Biomarkers Prev* 22.11 (2013), pp. 1984–93.

[221] P. Cohen et al. "Prostate-specific antigen (PSA) is an insulin-like growth factor binding protein-3 protease found in seminal plasma". In: *J Clin Endocrinol Metab* 75.4 (1992), pp. 1046–53.

[222] Y. C. Wong, X. H. Wang, and M. T. Ling. "Prostate development and carcinogenesis". In: *Int Rev Cytol* 227 (2003), pp. 65–130.

[223] S. J. McPherson et al. "Elevated androgens and prolactin in aromatase-deficient mice cause enlargement, but not malignancy, of the prostate gland". In: *Endocrinology* 142.6 (2001), pp. 2458–67.

[224] R. L. Noble. "The development of prostatic adenocarcinoma in Nb rats following prolonged sex hormone administration". In: *Cancer Res* 37.6 (1977), pp. 1929–33.

[225] H. Bonkhoff et al. "Estrogen receptor expression in prostate cancer and premalignant prostatic lesions". In: *Am J Pathol* 155.2 (1999), pp. 641–7.

[226] S. S. Taneja et al. "Prostate cancer diagnosis among men with isolated high-grade intraepithelial neoplasia enrolled onto a 3-year prospective phase III clinical trial of oral toremifene". In: *J Clin Oncol* 31.5 (2013), pp. 523–9.

[227] C. Selby. "Sex hormone binding globulin: origin, function and clinical significance". In: *Ann Clin Biochem* 27 (Pt 6) (1990), pp. 532–41.

[228] R. Simó et al. "Novel insights in SHBG regulation and clinical implications". In: *Trends Endocrinol Metab* 26.7 (2015), pp. 376–383.

[229] J. F. Dorgan et al. "Relationships of serum androgens and estrogens to prostate cancer risk: results from a prospective study in Finland". In: *Cancer Epidemiol Biomarkers Prev* 7.12 (1998), pp. 1069–74.

[230] H. B. Carter et al. "Longitudinal evaluation of serum androgen levels in men with and without prostate cancer". In: *Prostate* 27.1 (1995), pp. 25–31.

[231] J. K. Parsons et al. "Serum testosterone and the risk of prostate cancer: potential implications for testosterone therapy". In: *Cancer Epidemiol Biomarkers Prev* 14.9 (2005), pp. 2257–60.

[232] J. K. Morris et al. "Insulin-like growth factors and cancer: no role in screening. Evidence from the BUPA study and meta-analysis of prospective epidemiological studies". In: *Br J Cancer* 95.1 (2006), pp. 112–7.

[233] C. Chen et al. "Endogenous sex hormones and prostate cancer risk: a case-control study nested within the Carotene and Retinol Efficacy Trial". In: *Cancer Epidemiol Biomarkers Prev* 12.12 (2003), pp. 1410–6.

[234] C. J. Tsai et al. "Sex steroid hormones in young manhood and the risk of subsequent prostate cancer: a longitudinal study in African-Americans and Caucasians (United States)". In: *Cancer Causes Control* 17.10 (2006), pp. 1237–1244.

[235] C. Chen et al. "Prostate carcinoma incidence in relation to prediagnostic circulating levels of insulin-like growth factor I, insulin-like growth factor binding protein 3, and insulin". In: *Cancer* 103.1 (2005), pp. 76–84.

[236] J. V. Lacey et al. "Null Association between Insulin-like Growth Factors, Insulin-like Growth Factor-binding Proteins, and Prostate Cancer in a Prospective Study". In: *Cancer Epidemiol Biomarkers Prev* 10.10 (2001), pp. 1101–1102.

[237] R. C. Travis et al. "Serum androgens and prostate cancer among 643 cases and 643 controls in the European Prospective Investigation into Cancer and Nutrition". In: *Int J Cancer* 121.6 (2007), pp. 1331–8.

[238] Y. L. Low et al. "Polymorphisms in the CYP19 gene may affect the positive correlations between serum and urine phytoestrogen metabolites and plasma androgen concentrations in men". In: *J Nutr* 135.11 (2005), pp. 2680–2686.

[239] J. A. Janssen et al. "Circulating free insulin-like growth factor (IGF)-I, total IGF-I, and IGF binding protein-3 levels do not predict the future risk to develop prostate cancer: results of a case-control study involving 201 patients within a population-based screening with a 4-year interval". In: *J Clin Endocrinol Metab* 89.9 (2004), pp. 4391–6.

[240] R. Heikkila et al. "Serum testosterone and sex hormone-binding globulin concentrations and the risk of prostate carcinoma: a longitudinal study". In: *Cancer* 86.2 (1999), pp. 312–5.

[241] Z. Hyde et al. "Associations between testosterone levels and incident prostate, lung, and colorectal cancer. A population-based study". In: *Cancer Epidemiol Biomarkers Prev* 21.8 (2012), pp. 1319–1329.

[242] Y. X. Chan et al. "Higher dihydrotestosterone is associated with the incidence of lung cancer in older men". In: *Horm Cancer* 8.2 (2017), pp. 119–126.

[243] M. Manttari et al. "The Helsinki Heart Study: basic design and randomization procedure". In: *Eur Heart J* 8 Suppl I (1987), pp. 1–29.

[244] P. Stattin et al. "High levels of circulating testosterone are not associated with increased prostate cancer risk: a pooled prospective study". In: *Int J Cancer* 108.3 (2004), pp. 418–24.

[245] E. A. Platz et al. "Sex steroid hormones and the androgen receptor gene CAG repeat and subsequent risk of prostate cancer in the prostate-specific antigen era". In: *Cancer Epidemiol Biomarkers Prev* 14.5 (2005), pp. 1262–9.

[246] K. Ozasa et al. "Serum phytoestrogens and prostate cancer risk in a nested case-control study among Japanese men". In: *Cancer Sci* 95.1 (2004), pp. 65–71.

[247] A. M. Nomura et al. “Prediagnostic serum hormones and the risk of prostate cancer”. In: *Cancer Res* 48.12 (1988), pp. 3515–7.

[248] A. M. Nomura et al. “Serum androgens and prostate cancer”. In: *Cancer Epidemiol Biomarkers Prev* 5.8 (1996), pp. 621–5.

[249] N. Sawada et al. “Plasma testosterone and sex hormone-binding globulin concentrations and the risk of prostate cancer among Japanese men: a nested case-control study”. In: *Cancer Sci* 101.12 (2010), pp. 2652–7.

[250] L. J. Vatten et al. “Androgens in serum and the risk of prostate cancer: a nested case-control study from the Janus serum bank in Norway”. In: *Cancer Epidemiol Biomarkers Prev* 6.11 (1997), pp. 967–9.

[251] C. Schaefer et al. “IGF-I and prostate cancer”. In: *Science* 282.5387 (1998), pp. 199–199.

[252] G. Severi et al. “Circulating steroid hormones and the risk of prostate cancer”. In: *Cancer Epidemiol Biomarkers Prev* 15.1 (2006), pp. 86–91.

[253] G. Severi et al. “Circulating insulin-like growth factor-I and binding protein-3 and risk of prostate cancer”. In: *Cancer Epidemiol Biomarkers Prev* 15.6 (2006), pp. 1137–41.

[254] J. K. Gill et al. “Androgens, growth factors and risk of prostate cancer: the Multiethnic Cohort”. In: *Prostate* 70.8 (2010), pp. 906–915.

[255] B. A. Mohr et al. “Are serum hormones associated with the risk of prostate cancer? Prospective results from the Massachusetts Male Aging Study”. In: *Urology* 57.5 (2001), pp. 930–935.

[256] P. H. Gann et al. “Prospective study of sex hormone levels and risk of prostate cancer”. In: *J Natl Cancer Inst* 88.16 (1996), pp. 1118–26.

[257] G. L. Andriole et al. “Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up”. In: *J Natl Cancer Inst* 104.2 (2012), pp. 125–32.

[258] S. E. Oliver et al. “Screen-detected prostate cancer and the insulin-like growth factor axis: results of a population-based case-control study”. In: *Int J Cancer* 108.6 (2004), pp. 887–92.

[259] J. Donovan et al. “Prostate testing for cancer and treatment (ProtecT) feasibility study”. In: *Health Technol Assess* 7.14 (2003), pp. 1–88.

[260] M. A. Rowlands et al. “Circulating insulin-like growth factors and IGF-binding proteins in PSA-detected prostate cancer: the large case-control study ProtecT”. In: *Cancer Res* 72.2 (2012), pp. 503–15.

[261] R. L. Ferrini and E. Barrett-Connor. “Sex hormones and age: a cross-sectional study of testosterone and estradiol and their bioavailable fractions in community-dwelling men”. In: *Am J Epidemiol* 147.8 (1998), pp. 750–754.

[262] F. Meyer et al. “A prospective study of the insulin-like growth factor axis in relation with prostate cancer in the SU.VI.MAX trial”. In: *Cancer Epidemiol Biomarkers Prev* 14.9 (2005), pp. 2269–72.

[263] R. B. Hayes et al. "Etiologic and early marker studies in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial". In: *Control Clin Trials* 21.6 Suppl (2000), 349s–355s.

[264] NHS Digital. *Hospital Episode Statistics (HES)*. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics.

[265] *Hospital Episode Statistics (HES) analysis guide*. 2013. URL: http://docs.adrn.ac.uk/888040/mrdoc/pdf/888040_hes-analy-guide-apr13.pdf.

[266] National Audit Office. *Healthcare across the UK: A comparison of the NHS in England, Scotland, Wales and Northern Ireland*. 2012. URL: https://www.nao.org.uk/wp-content/uploads/2012/06/1213192.pdf.

[267] A. Herbert et al. "Data resource profile: Hospital Episode Statistics Admitted Patient Care (HES APC)". In: *Int J Epidemiol* 46.4 (2017), pp. 1093–1093i.

[268] NHS Digital. *National Clinical Coding Standards ICD-10 5th Edition (2017)*. Tech. rep. Clinical Classifications Service, 2017.

[269] The Health and Social Care Information Centre. *Replacement of the HES Patient ID (HESID)*. URL: https://www.cl.cam.ac.uk/~rja14/Papers/HESID_Replacement_Nov09.pdf.

[270] Health and Social Care Information Centre. *HES data dictionary: Admitted Patient Care*. URL: https://webarchive.nationalarchives.gov.uk/20180328130852tf_/http://content.digital.nhs.uk/media/19425/APC-DD-Final-Doc2/pdf/DD-APC-V7.pdf/.

[271] Office of the Deputy Prime Minister. *The English indices of deprivation 2004: summary (revised)*. URL: https://webarchive.nationalarchives.gov.uk/20100407220133/http://www.communities.gov.uk/documents/communities/pdf/131206.pdf.

[272] Health and Social Care Information Centre. *A guide to linked mortality data from Hospital Episode Statistics and the Office for National Statistics*. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/linked-hes-ons-mortality-data.

[273] G. Hagger-Johnson et al. "Data linkage errors in hospital administrative data when applying a pseudonymisation algorithm to paediatric intensive care records". In: *BMJ Open* 5.8 (2015), e008118.

[274] Office for National Statistics. *User guide to mortality statistics*. Tech. rep. 2018. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsj

[275] NHS Digital. *NHS Digital implement Type 2 patient opt-outs*. URL: https://mrc.ukri.org/documents/pdf/nhs-digital-and-type-2-patient-opt-outs-voct2016/.

[276] NHS Digital. *Opting out of sharing your confidential patient information.* URL: https://digital.nhs.uk/about-nhs-digital/our-work/keeping-patient-data-safe/how-we-look-after-your-health-and-care-information/your-information-choices/opting-out-of-sharing-your-confidential-patient-information.

[277] NHS Digital. *[MI] National Data Opt-out, March 2019.* URL: https://digital.nhs.uk/data-and-information/publications/statistical/national-data-opt-out/march-2019/ndop-mar19.

[278] N. E. Allen et al. "UK Biobank data: come and get it". In: *Sci Transl Med* 6.224 (2014), 224ed4–224ed4.

[279] UK Biobank. *UK Biobank: Protocol for a large-scale prospective epidemiological resource.* URL: https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf.

[280] P. Elliott and T. C. Peakman. "The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine". In: *Int J Epidemiol* 37.2 (2008), pp. 234–44.

[281] S.M. Sheard and J. Froggatt. *UK Biobank haematology data companion document.* URL: https://biobank.ctsu.ox.ac.uk/crystal/docs/haematology.pdf.

[282] UK Biobank. *Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications).* URL: http://biobank.ctsu.ox.ac.uk/showcase/docs/biomarker_issues.pdf.

[283] UK Biobank. *Cancer data: linkage from national cancer registries.* Tech. rep. 2013. URL: http://biobank.ndph.ox.ac.uk/showcase/docs/CancerLinkage.pdf.

[284] UK Biobank. *UK Biobank biomarker project companion document to accompany serum biomarker data.* URL: https://biobank.ndph.ox.ac.uk/showcase/docs/serum_biochemistry.pdf.

[285] UK Biobank. *Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource.* URL: http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UKBiobank_genotyping_QC_documentation-web.pdf.

[286] C. Bycroft et al. "The UK Biobank resource with deep phenotyping and genomic data". In: *Nature* 562.7726 (2018), pp. 203–209.

[287] UK Biobank. *UK Biobank axiom array content summary.* URL: http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UK-Biobank-Axiom-Array-Content-Summary-2014.pdf.

[288] UK Biobank. *24-hour dietary recall questionnaire.* URL: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/DietWebQ.pdf.

[289] K. E. Bradbury et al. "Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire". In: *J Nutr Sci* 7 (2018), e6–e6.

[290] R. Clarke et al. "Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies". In: *Am J Epidemiol* 150.4 (1999), pp. 341–53.

[291] UK Biobank. *Repeat assessment data*. URL: `https://biobank.ctsu.ox.ac.uk/~bbdatan/Repeat_assessment_doc_v1.0.pdf`.

[292] M. Maggi et al. "The burden of testosterone deficiency syndrome in adult men: economic and quality-of-life impact". In: *J Sex Med* 4.4 Pt 1 (2007), pp. 1056–69.

[293] E. Selvin et al. "Androgens and diabetes in men: results from the Third National Health and Nutrition Examination Survey (NHANES III)". In: *Diabetes Care* 30.2 (2007), pp. 234–8.

[294] A. Kaltenboeck et al. "The direct and indirect costs among U.S. privately insured employees with hypogonadism". In: *J Sex Med* 9.9 (2012), pp. 2438–47.

[295] R. K. Dubey et al. "Sex hormones and hypertension". In: *Cardiovasc Res* 53.3 (2002), pp. 688–708.

[296] C. Kim and J. B. Halter. "Endogenous sex hormones, metabolic syndrome, and diabetes in men and women". In: *Curr Cardiol Rep* 16.4 (2014), p. 467.

[297] S. Rohrmann et al. "Body fatness and sex steroid hormone concentrations in US men: results from NHANES III". In: *Cancer Causes Control* 22.8 (2011), pp. 1141–1151.

[298] Å Bjørnerem et al. "Endogenous sex hormones in relation to age, sex, lifestyle factors, and chronic diseases in a general population: The tromsø study". In: *J Clin Endocrinol Met* 89.12 (2004), pp. 6039–6047.

[299] S. M. Gapstur et al. "Serum androgen concentrations in young men: a longitudinal analysis of associations with age, obesity, and race. The CARDIA male hormone study". In: *Cancer Epidemiol Biomarkers Prev* 11.10 Pt 1 (2002), pp. 1041–7.

[300] J. E. Morley et al. "Longitudinal changes in testosterone, luteinizing hormone, and follicle-stimulating hormone in healthy older men". In: *Metabolism* 46.4 (1997), pp. 410–413.

[301] M. Muller et al. "Endogenous sex hormones in men aged 40-80 years". In: *Eur J Endocrinol* 149.6 (2003), pp. 583–9.

[302] W. Bartsch. "Interrelationships between sex hormone-binding globulin and testosterone, 5 alpha-dihydrotestosterone and oestradiol-17 beta in blood of normal men". In: *Maturitas* 2.2 (1980), pp. 109–18.

[303] F.C. McLean. "Application of the law of chemical equilibrium (law of mass action) to biological problems". In: *Physiol Rev* 18.4 (1938), pp. 495–523.

[304] F. Hu. *Obesity Epidemiology*. Vol. 9780195312911. Oxford University Press, 2008.

[305] S. M. Burton and S. T. Tiffany. "The effect of alcohol consumption on craving to smoke". In: *Addiction* 92.1 (1997), pp. 15–26.

[306] G. Sartorius et al. "Predictive accuracy and sources of variability in calculated free testosterone estimates". In: *Ann Clin Biochem* 46.Pt 2 (2009), pp. 137–43.

[307] K. J. Rothman. "No adjustments are needed for multiple comparisons". In: *Epidemiology* (1990), pp. 43–46.

[308] J. G. Lewis, R. Ghanadian, and G. D. Chisholm. "Serum 5α-dihydrotestosterone and testosterone changes with age in man". In: *Acta Endocrinol (Copenh)* 82.2 (1976), pp. 444–448.

[309] K. M. Pirke and P. Doerr. "Age related changes and interrelationships between plasma testosterone, oestradiol and testosterone-binding globulin in normal adult males". In: *Acta Endocrinol (Copenh)* 74.4 (1973), pp. 792–800.

[310] T. G. Travison et al. "The relative contributions of aging, health, and lifestyle factors to serum testosterone decline in men". In: *J Clin Endocrinol Metab* 92.2 (2007), pp. 549–55.

[311] H. A. Feldman et al. "Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts male aging study". In: *J Clin Endocrinol Metab* 87.2 (2002), pp. 589–98.

[312] Z. Shi et al. "Longitudinal changes in testosterone over five years in community-dwelling men". In: *J Clin Endocrinol Metab* 98.8 (2013), pp. 3289–97.

[313] E. M. Camacho et al. "Age-associated changes in hypothalamic-pituitary-testicular function in middle-aged and older men are modified by weight change and lifestyle factors: longitudinal results from the European Male Ageing Study". In: *Eur J Endocrinol* 168.3 (2013), pp. 445–55.

[314] F. C. Wu et al. "Hypothalamic-pituitary-testicular axis disruptions in older men are differentially linked to age and modifiable risk factors: the European Male Aging Study". In: *J Clin Endocrinol Metab* 93.7 (2008), pp. 2737–45.

[315] T.W. Kelsey et al. "A validated age-related normative model for male total testosterone shows increasing variance but no decline after age 40 years". In: *PLoS ONE* 9.10 (2014), e109346.

[316] M. Frost et al. "Similar reference intervals for total testosterone in healthy young and elderly men: results from the Odense Androgen Study". In: *Clin Endocrinol (Oxf)* 78.5 (2013), pp. 743–51.

[317] P. Marrama et al. "Circadian rhythm of testosterone and prolactin in the ageing". In: *Maturitas* 4.2 (1982), pp. 131–8.

[318] W. J. Bremner, M. V. Vitiello, and P. N. Prinz. "Loss of circadian rhythmicity in blood testosterone levels with aging in normal men". In: *J Clin Endocrinol Metab* 56.6 (1983), pp. 1278–81.

[319] P. Lecomte et al. "Sex differences in the control of sex-hormone-binding globulin in the elderly: role of insulin-like growth factor-I and insulin". In: *Eur J Endocrinol* 139.2 (1998), pp. 178–83.

[320] L. Parker et al. "Decreased adrenal androgen sensitivity to ACTH during aging". In: *Metabolism* 30.6 (1981), pp. 601–4.

[321] A. Belanger et al. "Changes in serum concentrations of conjugated and unconjugated steroids in 40- to 80-year-old men". In: *J Clin Endocrinol Metab* 79.4 (1994), pp. 1086–90.

[322] I. Morimoto et al. "Studies on the origin of androstanediol and androstanediol glucuronide in young and elderly men". In: *J Clin Endocrinol Metab* 52.4 (1981), pp. 772–8.

[323] A. Hautanen. "Synthesis and regulation of sex hormone-binding globulin in obesity". In: *Int J Obes Relat Metab Disord* 24.2 (2000), S64–70.

[324] S. R. Plymate et al. "Inhibition of sex hormone-binding globulin production in the human hepatoma (Hep G2) cell line by insulin and prolactin". In: *J Clin Endocrinol Metab* 67.3 (1988), pp. 460–4.

[325] P. G. Cohen. "Aromatase, adiposity, aging and disease. The hypogonadal-metabolic-atherogenic-disease and aging connection". In: *Med Hypotheses* 56.6 (2001), pp. 702–8.

[326] R. Pasquali and V. Vicennati. "Activity of the hypothalamic-pituitary-adrenal axis in different obesity phenotypes". In: *Int J Obes Relat Metab Disord* 24 Suppl 2 (2000), S47–9.

[327] A. Peter et al. "Relationships of circulating sex hormone-binding globulin with metabolic traits in humans". In: *Diabetes* 59.12 (2010), pp. 3167–73.

[328] V. Rochira, E. Kara, and C. Carani. "The endocrine role of estrogens on human male skeleton". In: *Int J Endocrinol* 2015 (2015), p. 15.

[329] R. J. Perry, C. Farquharson, and S. F. Ahmed. "The role of sex steroids in controlling pubertal growth". In: *Clin Endocrinol (Oxf)* 68.1 (2008), pp. 4–15.

[330] S. C. Manolagas, C. A. O'Brien, and M. Almeida. "The role of estrogen and androgen receptors in bone health and disease". In: *Nat Rev Endocrinol* 9.12 (2013), pp. 699–712.

[331] C. Dechanet et al. "Effects of cigarette smoking on reproduction". In: *Human Reproduction Update* 17.1 (2011), pp. 76–95.

[332] A. Mlynarcikova, M. Fickova, and S. Scsukova. "Ovarian intrafollicular processes as a target for cigarette smoke components and selected environmental reproductive disruptors". In: *Endocr Regul* 39.1 (2005), pp. 21–32.

[333] W. De Ronde et al. "Serum levels of sex hormone-binding globulin (SHBG) are not associated with lower levels of non-SHBG-bound testosterone in male newborns and healthy adult men". In: *Clin Endocrinol (Oxf)* 62.4 (2005), pp. 498–503.

[334] K. M. English et al. “Effect of cigarette smoking on levels of bioavailable testosterone in healthy men”. In: *Clin Sci (Lond)* 100.6 (2001), pp. 661–5.

[335] J. Svartberg and R. Jorde. “Endogenous testosterone levels and smoking in men. The fifth Tromsø study”. In: *Int J Androl* 30.3 (2007), pp. 137–143.

[336] S. Kiechl et al. “No association between dehydroepiandrosterone sulfate and development of atherosclerosis in a prospective population study (Bruneck Study)”. In: *Arterioscler Thromb Vasc Biol* 20.4 (2000), pp. 1094–100.

[337] G. Ravaglia et al. “Dehydroepiandrosterone-sulfate serum levels and common age-related diseases: results from a cross-sectional Italian study of a general elderly population”. In: *Exp Gerontol* 37.5 (2002), pp. 701–12.

[338] J. Frias et al. “Effects of acute alcohol intoxication on pituitary-gonadal axis hormones, pituitary-adrenal axis hormones, β-endorphin and prolactin in human adults of both sexes”. In: *Alcohol Alcohol* 37.2 (2002), pp. 169–173.

[339] O. Ukkola et al. “Age, body mass index, race and other determinants of steroid hormone variability: the HERITAGE Family Study”. In: *Eur J Endocrinol* 145.1 (2001), pp. 1–9.

[340] A. H. Wu et al. “Serum androgens and sex hormone-binding globulins in relation to lifestyle factors in older African-American, white, and Asian men in the United States and Canada”. In: *Cancer Epidemiol Biomarkers Prev* 4.7 (1995), pp. 735–41.

[341] D. Sparrow, R. Bosse, and J. W. Rowe. “The influence of age, alcohol consumption, and body build on gonadal function in men”. In: *J Clin Endocrinol Metab* 51.3 (1980), pp. 508–12.

[342] D. P. Lookingbill et al. “Clinical and biochemical parameters of androgen action in normal healthy Caucasian versus Chinese subjects”. In: *J Clin Endocrinol Metab* 72.6 (1991), pp. 1242–8.

[343] A. H. Wu et al. “Lifestyle determinants of 5alpha-reductase metabolites in older African-American, white, and Asian-American men”. In: *Cancer Epidemiol Biomarkers Prev* 10.5 (2001), pp. 533–8.

[344] D. J. Handelsman and L. Wartofsky. “Requirement for mass spectrometry sex steroid assays in the Journal of Clinical Endocrinology and Metabolism”. In: *J Clin Endocrinol Metab* 98.10 (2013), pp. 3971–3.

[345] W. Rosner et al. “Position statement: Utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement”. In: *J Clin Endocrinol Metab* 92.2 (2007), pp. 405–13.

[346] S. Rinaldi et al. “Validity of free testosterone and free estradiol determinations in serum samples from postmenopausal women by theoretical calculations”. In: *Cancer Epidemiol Biomarkers Prev* 11.10 (2002), pp. 1065–1071.

[347] L. P. Ly and D. J. Handelsman. “Empirical estimation of free testosterone from testosterone and sex hormone-binding globulin immunoassays”. In: *Eur J Endocrinol* 152.3 (2005), pp. 471–8.

[348] C. K. Ho et al. “Calculated free testosterone in men: comparison of four equations and with free androgen index”. In: *Ann Clin Biochem* 43.Pt 5 (2006), pp. 389–97.

[349] A. Vermeulen, L. Verdonck, and J. M. Kaufman. “A critical evaluation of simple methods for the estimation of free testosterone in serum”. In: *J Clin Endocrinol Metab* 84.10 (1999), pp. 3666–72.

[350] L. P. Ly et al. “Accuracy of calculated free testosterone formulae in men”. In: *Clin Endocrinol (Oxf)* 73.3 (2010), pp. 382–388.

[351] T. Fiers et al. “Reassessing free-testosterone calculation by liquid chromatography–tandem mass spectrometry direct equilibrium dialysis”. In: *J Clin Endocrinol Metab* 103.6 (2018), pp. 2167–2174.

[352] W. de Ronde et al. “Calculation of bioavailable and free testosterone in men: a comparison of 5 published algorithms”. In: *Clin Chem* 52.9 (2006), pp. 1777–84.

[353] V. N. Hawkins et al. “Effect of exercise on serum sex hormones in men: a 12-month randomized clinical trial”. In: *Med Sci Sports Exerc* 40.2 (2008), pp. 223–233.

[354] N. E. Allen et al. “Lifestyle and nutritional determinants of bioavailable androgens and related hormones in British men”. In: *Cancer Causes Control* 13.4 (2002), pp. 353–63.

[355] A. Vermeulen and G. Verdonck. “Representativeness of a single point plasma testosterone level for the long term hormonal milieu in men”. In: *J Clin Endocrinol Metab* 74.4 (1992), pp. 939–42.

[356] M. E. Jones et al. “Reproducibility of estradiol and testosterone levels in postmenopausal women over 5 Years: Results from the breakthrough generations study”. In: *Am J Epidemiol* 179.9 (2014), pp. 1128–1133.

[357] A. N. Tint et al. “Association of sex hormone-binding globulin and free testosterone with mortality in men with type 2 diabetes mellitus”. In: *Eur J Endocrinol* 174.1 (2016), p. 59.

[358] P. Muti et al. “Reliability of serum hormones in premenopausal and postmenopausal women over a one-year period”. In: *Cancer Epidemiol Biomarkers Prev* 5.11 (1996), pp. 917–22.

[359] J. M. Holly and C. M. Perks. “Insulin-like growth factor physiology: what we have learned from human studies”. In: *Endocrinol Metab Clin North Am* 41.2 (2012), pp. 249–63, v.

[360] V. G. Kaklamani et al. “Age, sex, and smoking are predictors of circulating insulin-like growth factor 1 and insulin-like growth factor-binding protein 3”. In: *J Clin Oncol* 17.3 (1999), pp. 813–7.

[361] N. E. Allen et al. “Lifestyle determinants of serum insulin-like growth-factor-I (IGF-I), C-peptide and hormone binding protein levels in British women”. In: *Cancer Causes Control* 14.1 (2003), pp. 65–74.

[362] C. J. Stein and G. A. Colditz. "Modifiable risk factors for cancer". In: *Br J Cancer* 90.2 (2004), pp. 299–303.

[363] A. J. Sasco, M. B. Secretan, and K. Straif. "Tobacco smoking and cancer: a brief review of recent epidemiological evidence". In: *Lung Cancer* 45 (2004), S3–S9.

[364] S. L. Siegel et al. "Cancer statistics for Hispanics/Latinos, 2015". In: *CA Cancer J Clin* 65.6 (2015), pp. 457–480.

[365] American Cancer Society. "Cancer facts and figures for African Americans 2016-2018". In: *Atlanta: American Cancer Society* (2016). URL: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/cancer-facts-and-figures-for-african-americans/cancer-facts-and-figures-for-african-americans-2016-2018.pdf.

[366] F. L. Crowe et al. "The association between diet and serum concentrations of IGF-I, IGFBP-1, IGFBP-2, and IGFBP-3 in the European Prospective Investigation into Cancer and Nutrition". In: *Cancer Epidemiol Biomarkers Prev* 18.5 (2009), pp. 1333–40.

[367] J. M. Faupel-Badger et al. "Anthropometric correlates of insulin-like growth factor 1 (IGF-1) and IGF binding protein-3 (IGFBP-3) levels by race/ethnicity and gender". In: *Ann Epidemiol* 19.12 (2009), pp. 841–9.

[368] E. A. Platz et al. "Racial variation in insulin-like growth factor-1 and binding protein-3 concentrations in middle-aged men". In: *Cancer Epidemiol Biomarkers Prev* 8.12 (1999), pp. 1107–10.

[369] J. H. Fowke et al. "Racial differences in the association between body mass index and serum IGF1, IGF2, and IGFBP3". In: *Endocr Relat Cancer* 17.1 (2010), pp. 51–60.

[370] K. Woodson et al. "Serum insulin-like growth factor I: tumor marker or etiologic factor? A prospective study of prostate cancer among Finnish men". In: *Cancer Res* 63.14 (2003), pp. 3991–4.

[371] S. M. Harman et al. "Serum levels of insulin-like growth factor I (IGF-I), IGF-II, IGF-binding protein-3, and prostate-specific antigen as predictors of clinical prostate cancer". In: *J Clin Endocrinol Metab* 85.11 (2000), pp. 4258–4265.

[372] B. J. van den Berg, R. E. Christianson, and F. W. Oechsli. "The California Child Health and Development Studies of the School of Public Health, University of California at Berkeley". In: *Paediatr Perinat Epidemiol* 2.3 (1988), pp. 265–82.

[373] P. E. Norman et al. "Cohort Profile: The Health In Men Study (HIMS)". In: *Int J Epidemiol* 38.1 (2009), pp. 48–52.

[374] B. B. Yeap et al. "IGF1 and its binding proteins 3 and 1 are differentially associated with metabolic syndrome in older men". In: *Eur J Endocrinol* 162.2 (2010), pp. 249–257.

[375] K. Nimptsch et al. “Plasma insulin-like growth factor 1 is positively associated with low-grade prostate cancer in the Health Professionals Follow-up Study 1993-2004”. In: *Int J Cancer* 128.3 (2011), pp. 660–7.

[376] E. A. Platz et al. “Plasma insulin-like growth factor-1 and binding protein-3 and subsequent risk of prostate cancer in the PSA era”. In: *Cancer Causes Control* 16.3 (2005), pp. 255–62.

[377] P. Stattin et al. “Plasma insulin-like growth factor-I, insulin-like growth factor-binding proteins, and prostate cancer risk: a prospective study”. In: *J Natl Cancer Inst* 92.23 (2000), pp. 1910–7.

[378] M. L. Neuhouser et al. “Insulin-like growth factors and insulin-like growth factor-binding proteins and prostate cancer risk: results from the prostate cancer prevention trial”. In: *Cancer Prev Res (Phila)* 6.2 (2013), pp. 91–9.

[379] J. M. Chan et al. “Plasma insulin-like growth factor-I and prostate cancer risk: a prospective study”. In: *Science* 279.5350 (1998), pp. 563–6.

[380] J. M. Chan et al. “Insulin-like growth factor-I (IGF-I) and IGF binding protein-3 as predictors of advanced-stage prostate cancer”. In: *J Natl Cancer Inst* 94.14 (2002), pp. 1099–1106.

[381] L. A. Mucci et al. “Plasma levels of acid-labile subunit, free insulin-like growth factor-I, and prostate cancer risk: a prospective study”. In: *Cancer Epidemiol Biomarkers Prev* 19.2 (2010), pp. 484–91.

[382] J. M. Weiss et al. “IGF-1 and IGFBP-3: Risk of prostate cancer among men in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial”. In: *Int J Cancer* 121.10 (2007), pp. 2267–73.

[383] A. Soliman et al. “Advances in pubertal growth and factors influencing it: Can we increase pubertal growth?” In: *Indian J Endocrinol Metab* 18.Suppl 1 (2014), S53–S62.

[384] A. M. Cotterill, J. M. Holly, and J. A. Wass. “The regulation of insulin-like growth factor binding protein (IGFBP)-1 during prolonged fasting”. In: *Clin Endocrinol (Oxf)* 39.3 (1993), pp. 357–62.

[385] H. S. Kang et al. “Regulation of IGFBP-2 expression during fasting”. In: *Biochem J.* 467.Pt 3 (2015), pp. 453–460.

[386] L. A. Frohman, T. R. Downs, and P. Chomczynski. “Regulation of growth hormone secretion”. In: *Front Neuroendocrinol* 13.4 (1992), pp. 344–405.

[387] H. Olivecrona et al. “Acute and short-term effects of growth hormone on insulin-like growth factors and their binding proteins: serum levels and hepatic messenger ribonucleic acid responses in Humans”. In: *J Clin Endocrinol Metab* 84.2 (1999), pp. 553–560.

[388] J. Brouwer-Visser and G. S. Huang. “IGF2 signaling and regulation in cancer”. In: *Cytokine Growth Factor Rev* 26.3 (2015), pp. 371–377.

[389] A. M. Suikkari et al. “Dose-response characteristics for suppression of low molecular weight plasma insulin-like growth factor-binding protein by insulin”. In: *J Clin Endocrinol Metab* 68.1 (1989), pp. 135–40.

[390] K. Brismar et al. “Effect of insulin on the hepatic production of insulin-like growth factor-binding protein-1 (IGFBP-1), IGFBP-3, and IGF-I in insulin-dependent diabetes”. In: *J Clin Endocrinol Metab* 79.3 (1994), pp. 872–8.

[391] M. Shin et al. “Recent insights into insulin-like growth factor binding protein 2 transcriptional regulation”. In: *Endocrinol Metab (Seoul)* 32.1 (2017), pp. 11–17.

[392] P. Cohen et al. “Insulin-like growth factors (IGFs): implications for aging”. In: *Psychoneuroendocrinology* 17.4 (1992), pp. 335–42.

[393] E. Corpas, S. M. Harman, and M. R. Blackman. “Human growth hormone and human aging”. In: *Endocr Rev* 14.1 (1993), pp. 20–39.

[394] R. I. Fink et al. “Mechanisms of insulin resistance in aging”. In: *J Clin Invest* 71.6 (1983), pp. 1523–1535.

[395] H. A. Heald et al. “Close relation of fasting insulin-like growth factor binding protein-1 (IGFBP-1) with glucose tolerance and cardiovascular risk in two populations”. In: *Diabetologia* 44.3 (2001), pp. 333–339.

[396] M. S. Lewitt, M. S. Dent, and K. Hall. “The insulin-like growth factor system in obesity, insulin resistance and type 2 diabetes mellitus”. In: *J Clin Med* 3.4 (2014), pp. 1561–1574.

[397] P. D. Lee et al. “Insulin-like growth factor binding protein-1: recent findings and new directions”. In: *Proc Soc Exp Biol Med* 216.3 (1997), pp. 319–57.

[398] M. S. Sandhu et al. “Circulating concentrations of insulin-like growth factor-I and development of glucose intolerance: a prospective observational study”. In: *Lancet* 359.9319 (2002), pp. 1740–5.

[399] U. Petersson et al. “Low levels of insulin-like growth-factor-binding protein-1 (IGFBP-1) are prospectively associated with the incidence of type 2 diabetes and impaired glucose tolerance (IGT): the Soderakra Cardiovascular Risk Factor Study”. In: *Diabetes Metab* 35.3 (2009), pp. 198–205.

[400] E. Giovannucci et al. “Diabetes and cancer: A consensus report”. In: *Diabetes Care* 33.7 (2010), pp. 1674–1685.

[401] A. G. Renehan et al. “The effect of cigarette smoking use and cessation on serum insulin-like growth factors”. In: *Br J Cancer* 91.8 (2004), pp. 1525–31.

[402] S. Teramukai et al. “Anthropometric and behavioral correlates of insulin-like growth factor I and insulin-like growth factor binding protein 3 in middle-aged Japanese men”. In: *Am J Epidemiol* 156.4 (2002), pp. 344–8.

[403] J. A. Janssen et al. “Serum total IGF-I, free IGF-I, and IGFBP-1 levels in an elderly population: Relation to cardiovascular risk factors and disease”. In: *Arterioscler Thromb Vasc Biol* 18.2 (1998), pp. 277–282.

[404] M. Berry et al. "Proteomics analysis reveals IGFBP2 as a candidate diagnostic biomarker for heart failure". In: *IJC Metabolic Endocrine* 6 (2015), pp. 5–12.

[405] M. Huncharek et al. "Smoking as a risk factor for prostate cancer: A meta-analysis of 24 prospective cohort studies". In: *Am J Public Health* 100.4 (2010), pp. 693–701.

[406] C. De Nunzio et al. "Smoking and prostate cancer: A systematic review". In: *Eur Urol Focus* 1.1 (2015), pp. 28–38.

[407] M. J. Ronis et al. "Alcohol-induced disruption of endocrine signaling". In: *Alcohol Clin Exp Res* 31.8 (2007), pp. 1269–85.

[408] J. Liu et al. "Pericentral hepatocytes produce insulin-like growth factor-2 to promote liver regeneration during selected injuries in mice". In: *Hepatology* 66.6 (2017), pp. 2002–2015.

[409] S. Rojdmark et al. "Insulin-like growth factor (IGF)-1 and IGF-binding protein-1 concentrations in serum of normal subjects after alcohol ingestion: evidence for decreased IGF-1 bioavailability". In: *Clin Endocrinol (Oxf)* 52.3 (2000), pp. 313–8.

[410] M. Knip et al. "Ethanol induces a paradoxical simultaneous increase in circulating concentrations of insulin-like growth factor binding protein-1 and insulin". In: *Metabolism* 44.10 (1995), pp. 1356–1359.

[411] M. K. Downer et al. "Alcohol intake and risk of lethal prostate cancer in the Health Professionals Follow-Up Study". In: *J Clin Oncol* 37.17 (2019), pp. 1499–1511.

[412] D. Berrigan et al. "Race/ethnic variation in serum levels of IGF-I and IGFBP-3 in US adults". In: *Growth Horm IGF Res* 19.2 (2009), pp. 146–55.

[413] S. Turan et al. "The effect of economic status on height, insulin-like growth factor (IGF)-I and IGF binding protein-3 concentrations in healthy Turkish children". In: *Eur J Clin Nutr* 61 (2006), p. 752.

[414] G. A. Aguirre et al. "Insulin-like growth factor-1 deficiency and metabolic syndrome". In: *J Transl Med* 14.1 (2016), p. 3.

[415] J. Frystyk, P. Freda, and D. R. Clemmons. "The current status of IGF-I assays–a 2009 update". In: *Growth Horm IGF Res* 20.1 (2010), pp. 8–18.

[416] D. R. Clemmons. "Consensus statement on the standardization and evaluation of growth hormone and insulin-like growth factor assays". In: *Clin Chem* 57.4 (2011), pp. 555–559.

[417] R. Kaaks et al. "Serum C-peptide, insulin-like growth factor (IGF)-I, IGF-binding proteins, and colorectal cancer risk in women". In: *J Natl Cancer Inst* 92.19 (2000), pp. 1592–1600.

[418] K. J. Hunt et al. "Insulin-like growth factor II and colorectal cancer risk in women". In: *Cancer Epidemiol Biomarkers Prev* 11.9 (2002), p. 901.

[419] A. Eliakim et al. "Increased physical activity and the growth hormone-IGF-I axis in adolescent males". In: *Am J Physiol* 275.1 Pt 2 (1998), R308–14.

[420] A. H. Heald et al. "The influence of dietary intake on the insulin-like growth factor (IGF) system across three ethnic groups: a population-based study". In: *Public Health Nutr* 6.2 (2003), pp. 175–80.

[421] S. Teppala and A. Shankar. "Association between serum IGF-1 and diabetes among U.S. adults". In: *Diabetes Care* 33.10 (2010), pp. 2257–2259.

[422] S. Teppala, A. Shankar, and C. Sabanayagam. "Association between IGF-1 and chronic kidney disease among US adults". In: *Clin Exp Nephrol* 14.5 (2010), pp. 440–4.

[423] E. Giovannucci et al. "Nutritional predictors of insulin-like growth factor I and their relationships to cancer in men". In: *Cancer Epidemiol Biomarkers Prev* 12.2 (2003), pp. 84–9.

[424] B. S. Taylor et al. "Integrative genomic profiling of human prostate cancer". In: *Cancer Cell* 18.1 (2010), pp. 11–22.

[425] R. A. Eeles et al. "Multiple newly identified loci associated with prostate cancer susceptibility". In: *Nat Genet* 40.3 (2008), pp. 316–21.

[426] M. Pollard and P. H. Luckert. "Autochthonous prostate adenocarcinomas in Lobund-Wistar rats: a model system". In: *Prostate* 11.3 (1987), pp. 219–27.

[427] T. M. van der Sluis et al. "Intraprostatic testosterone and dihydrotestosterone. Part II: concentrations after androgen hormonal manipulation in men with benign prostatic hyperplasia and prostate cancer". In: *BJU Int* 109.2 (2012), pp. 183–188.

[428] T. J. Key et al. "Pooling biomarker data from different studies of disease risk, with a focus on endogenous hormones". In: *Cancer Epidemiol Biomarkers Prev* 19.4 (2010), pp. 960–965.

[429] A.W. Hsing et al. "Reproducibility of serum sex steroid assays in men by RIA and mass spectrometry". In: *Cancer Epidemiol Biomarkers Prev* 16.5 (2007), pp. 1004–1008.

[430] M. Plummer. "Improved estimates of floating absolute risk". In: *Stat Med* 23.1 (2004), pp. 93–104.

[431] R. L. Muller et al. "Serum testosterone and dihydrotestosterone and prostate cancer risk in the placebo arm of the Reduction by Dutasteride of Prostate Cancer Events Trial". In: *Euro Urol* 62.5 (2012), pp. 757–764.

[432] G. Schatzl et al. "High-grade prostate cancer is associated with low serum testosterone levels". In: *Prostate* 47.1 (2001), pp. 52–8.

[433] M. A. Hoffman, W. C. DeWolf, and A. Morgentaler. "Is low serum free testosterone a marker for high grade prostate cancer?" In: *J Urol* 163.3 (2000), pp. 824–7.

[434] G. Bayar et al. "Low free and bioavailable testosterone levels may predict pathologically-proven high-risk prostate cancer: a prospective, clinical trial". In: *Turk J Urol* 43.3 (2017), pp. 289–296.

[435] K. L. Penney et al. “Gleason grade progression is uncommon”. In: *Cancer Res* 73.16 (2013), pp. 5163–5168.

[436] G. R. Cunha et al. “The endocrinology and developmental biology of the prostate”. In: *Endocr Rev* 8.3 (1987), pp. 338–62.

[437] A. Briganti et al. “Prostate volume and adverse prostate cancer features: fact not artifact”. In: *Eur J Cancer* 43.18 (2007), pp. 2669–77.

[438] Y. C. Cohen et al. “Detection bias due to the effect of finasteride on prostate volume: a modeling approach for analysis of the Prostate Cancer Prevention Trial”. In: *J Natl Cancer Inst* 99.18 (2007), pp. 1366–74.

[439] M. S. Lucia et al. “Finasteride and high-grade prostate cancer in the Prostate Cancer Prevention Trial”. In: *J Natl Cancer Inst* 99.18 (2007), pp. 1375–83.

[440] F. Civantos, M. S. Soloway, and J. E. Pinto. “Histopathological effects of androgen deprivation in prostatic cancer”. In: *Semin Urol Oncol* 14.2 Suppl 2 (1996), pp. 22–31.

[441] V. E. Reuter. “Pathological changes in benign and malignant prostatic tissue following androgen deprivation therapy”. In: *Urology* 49.3A Suppl (1997), pp. 16–22.

[442] I. M. Thompson et al. “Effect of finasteride on the sensitivity of PSA for detecting prostate cancer”. In: *J Natl Cancer Inst* 98.16 (2006), pp. 1128–1133.

[443] F. A. Faisal and J. J. Tosoian. “The continuing debate of 5-alpha reductase inhibitors and prostate cancer risk”. In: *Transl Cancer Res* 7.6 (2018), pp. 1344–1346.

[444] J. M. Lacy and N. Kyprianou. “A tale of two trials: The impact of 5α-reductase inhibition on prostate cancer (Review)”. In: *Oncol Lett* 8.4 (2014), pp. 1391–1396.

[445] J. Kim and G. A. Coetzee. “Prostate specific antigen gene regulation by androgen receptor”. In: *J Cell Biochem* 93.2 (2004), pp. 233–41.

[446] J. M. Legler et al. “The role of prostate-specific antigen (PSA) testing patterns in the recent prostate cancer incidence decline in the United States”. In: *Cancer Causes Control* 9.5 (1998), pp. 519–27.

[447] J. A. Osuna C et al. “Relationship between BMI, total testosterone, sex hormone-binding-globulin, leptin, insulin and insulin resistance in obese men”. In: *Arch Androl* 52.5 (2006), pp. 355–361.

[448] P. Dandona and S. Dhindsa. “Update: hypogonadotropic hypogonadism in type 2 diabetes and obesity”. In: *J Clin Endocrinol Metab* 96.9 (2011), pp. 2643–2651.

[449] M. Grossmann. “Low testosterone in men with type 2 diabetes: significance and treatment”. In: *J Clin Endocrinol Metab* 96.8 (2011), pp. 2341–2353.

[450] C. Rodriguez et al. "Body mass index, weight change, and risk of prostate cancer in the Cancer Prevention Study II Nutrition Cohort". In: *Cancer Epidemiol Biomarkers Prev* 16.1 (2007), pp. 63–69.

[451] J. Baillargeon et al. "The association of body mass index and prostate-specific antigen in a population-based study". In: *Cancer* 103.5 (2005), pp. 1092–1095.

[452] C. M. Tangen et al. "Biases in recommendations for and acceptance of prostate biopsy significantly affect assessment of prostate cancer risk factors: results from two large randomized clinical trials". In: *J Clin Oncol* 34.36 (2016), pp. 4338–4344.

[453] C. Wang et al. "Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography-tandem mass spectrometry". In: *J Clin Endocrinol Metab* 89.2 (2004), pp. 534–43.

[454] NHS Digital. *Provisional Monthly Hospital Episode Statistics for Admitted Patient Care, Outpatients and Accident and Emergency data, April 2016 to March 2017*. URL: `https://files.digital.nhs.uk/publication/c/o/m13_hesf_monthly_report.pdf`.

[455] A. F. Daly et al. "High prevalence of pituitary adenomas: a cross-sectional study in the province of Liege, Belgium". In: *J Clin Endocrinol Metab* 91.12 (2006), pp. 4769–75.

[456] A. Giustina et al. "A consensus on criteria for cure of acromegaly". In: *J Clin Endocrinol Metab* 95.7 (2010), pp. 3141–3148.

[457] S. Zhang et al. "Body mass index and insulin-like growth factor 1 as risk factors for discordant growth hormone and insulin-like growth factor 1 levels following pituitary surgery in acromegaly". In: *J Formos Med Assoc* 117.1 (2018), pp. 34–41.

[458] S. Melmed et al. "A Consensus Statement on acromegaly therapeutic outcomes". In: *Nat Rev Endocrinol* 14.9 (2018), pp. 552–561.

[459] J. Dal et al. "Cancer incidence in patients with acromegaly: a cohort study and meta-analysis of the literature". In: *J Clin Endocrinol Metab* 103.6 (2018), pp. 2182–2188.

[460] K. A. Groth et al. "Klinefelter syndrome—a clinical update". In: *J Clin Endocrinol Metab* 98.1 (2013), pp. 20–30.

[461] A. Bojesen, S. Juul, and C. H. Gravholt. "Prenatal and postnatal prevalence of Klinefelter syndrome: a national registry study". In: *J Clin Endocrinol Metab* 88.2 (2003), pp. 622–6.

[462] L. Abramsky and J. Chapple. "47,XXY (Klinefelter syndrome) and 47,XYY: estimated rates of and indication for postnatal diagnosis with implications for prenatal counselling". In: *Prenat Diagn* 17.4 (1997), pp. 363–8.

[463] A. J. Swerdlow et al. "Cancer incidence and mortality in men with Klinefelter Syndrome: a cohort study". In: *J Natl Cancer Inst* 97.16 (2005), pp. 1204–1210.

[464] A. S Herlihy et al. "The prevalence and diagnosis rates of Klinefelter syndrome: an Australian comparison". In: *Med J Aust* 194.1 (2011), pp. 24–28.

[465] F. Lanfranco et al. "Klinefelter's syndrome". In: *Lancet* 364.9430 (2004), pp. 273–283.

[466] N. Pacenza et al. "Clinical presentation of Klinefelter's syndrome: Differences according to age". In: *Int J Endocrinol* 2012 (2012), p. 6.

[467] V. De Sanctis et al. "Klinefelter syndrome and cancer: from childhood to adulthood". In: *Pediatr Endocrinol Rev* 11.1 (2013), pp. 44–50.

[468] G. R. Dohle et al. "EAU guidelines on male hypogonadism". In: *European Association of Urology, Arnhem, The Netherlands, accessed Mar* (2018). URL: `https://uroweb.org/wp-content/uploads/EAU-Guidelines-Male-Hypogonadism-2016-1.pdf`.

[469] H. Hasle et al. "Cancer incidence in men with Klinefelter syndrome". In: *Br J Cancer* 71.2 (1995), pp. 416–420.

[470] J. Jianguang et al. "Risk of solid tumors and hematological malignancy in persons with Turner and Klinefelter syndromes: A national cohort study". In: *Int J Cancer* 139.4 (2016), pp. 754–758.

[471] C. R. Kannan. "Testicular hypofunction". In: *Essential endocrinology: A primer for nonspecialists*. Ed. by C. R. Kannan. Boston, MA: Springer US, 1986, pp. 307–322.

[472] S. Basaria. "Male hypogonadism". In: *Lancet* 383.9924 (2014), pp. 1250–1263.

[473] C. Carnegie. "Diagnosis of hypogonadism: clinical assessments and laboratory tests". In: *Rev Urol* 6 Suppl 6.Suppl 6 (2004), S3–S8.

[474] A. Tajar et al. "Characteristics of secondary, primary, and compensated hypogonadism in aging men: evidence from the European Male Ageing Study". In: *J Clin Endocrinol Metab* 95.4 (2010), pp. 1810–8.

[475] M. Regal et al. "Prevalence and incidence of hypopituitarism in an adult Caucasian population in northwestern Spain". In: *Clin Endocrinol (Oxf)* 55.6 (2001), pp. 735–740.

[476] P. Ascoli and F. Cavagnini. "Hypopituitarism". In: *Pituitary* 9.4 (2006), pp. 335–42.

[477] M. O. van Aken and S. W. Lamberts. "Diagnosis and treatment of hypopituitarism: an update". In: *Pituitary* 8.3-4 (2005), pp. 183–91.

[478] WHO. "International statistical classification of diseases and related health problems. 10th revision." In: (2010). URL: `https://www.who.int/classifications/icd/en/`.

[479] Department for Communities and Local Government. *The English index of multiple deprivation (IMD) 2015–guidance*. 2015. URL: `https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015`.

[480] C. A. Salinas et al. "Prostate cancer in young men: an important clinical entity". In: *Nat Rev Urol* 11.6 (2014), pp. 317–323.

[481] E. L. Rhoden and A. Morgentaler. "Risks of testosterone-replacement therapy and recommendations for monitoring". In: *N Engl J Med* 350.5 (2004), pp. 482–92.

[482] G. Lugo, L. Pena, and F. Cordido. "Clinical manifestations and diagnosis of acromegaly". In: *Int J Endocrinol* 2012 (2012), p. 540398.

[483] M. Arosio et al. "Diagnostic value of the acid-labile subunit in acromegaly: evaluation in comparison with insulin-like growth factor (IGF) I, and IGF-binding protein-1, -2, and -3". In: *J Clin Endocrinol Metab* 86.3 (2001), pp. 1091–1098.

[484] S. Melmed. "Acromegaly". In: *N Engl J Med* 355.24 (2006), pp. 2558–2573.

[485] M. Paja et al. "Hypothalamic-pituitary dysfunction in patients with craniopharyngioma". In: *Clin Endocrinol (Oxf)* 42.5 (1995), pp. 467–73.

[486] J. DiGiovanni et al. "Deregulated expression of insulin-like growth factor 1 in prostate epithelium leads to neoplasia in transgenic mice". In: *Pro Natl Acad Sci USA* 97.7 (2000), pp. 3455–3460.

[487] C. A. Heinlein and C. Chang. "Androgen receptor in prostate cancer". In: *Endocr Rev* 25.2 (2004), pp. 276–308.

[488] R. L. Ernani and A. Morgentaler. "Testosterone replacement therapy in hypogonadal men at high risk for prostate cancer: Results of 1 year of treatment in men with prostatic intraepithelial neoplasia". In: *J Urol* 170.6 (2003), pp. 2348–2351.

[489] P. J. Snyder et al. "Lessons from the testosterone trials". In: *Endocrine Reviews* 39.3 (2018), pp. 369–386.

[490] R. Boada et al. "The cognitive phenotype in Klinefelter syndrome: a review of the literature including genetic and hormonal factors". In: *Dev Disabil Res Rev* 15.4 (2009), pp. 284–294.

[491] A. Verri et al. "Klinefelter's syndrome and psychoneurologic function". In: *Mol Hum Reprod* 16.6 (2010), pp. 425–433.

[492] L. Katznelson et al. "Acromegaly: An Endocrine Society clinical practice guideline". In: *J Clin Endocrinol Metab* 99.11 (2014), pp. 3933–3951.

[493] G. Draisma et al. "Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer". In: *J Natl Cancer Inst* 95.12 (2003), pp. 868–78.

[494] J. L. Jahn, E. L. Giovannucci, and M. J. Stampfer. "The high prevalence of undiagnosed prostate cancer at autopsy: implications for epidemiology and treatment of prostate cancer in the prostate-specific antigen-era". In: *Int J Cancer* 137.12 (2015), pp. 2795–2802.

[495] R. Zahr and M. Fleseriu. "Updates in diagnosis and treatment of acromegaly". In: *Eur Endocrinol* 14.2 (2018), pp. 57–61.

[496] Peter J. Snyder et al. "Effects of testosterone treatment in older men". In: *N Engl J Med* 374.7 (2016), pp. 611–624. URL: https://doi.org/10.1056/NEJMoa1506119.

[497] E. M. Burns et al. "Systematic review of discharge coding accuracy". In: *J Public Health (Oxf)* 34.1 (2012), pp. 138–48.

[498] P. Valent et al. "Normal and pathological erythropoiesis in adults: from gene regulation to targeted treatment concepts". In: *Haematologica* 103.10 (2018), pp. 1593–1603.

[499] A. D. Coviello et al. "Effects of graded doses of testosterone on erythropoiesis in healthy young and older men". In: *J Clin Endocrinol Metab* 93.3 (2008), pp. 914–919.

[500] L. Ferrucci et al. "Low testosterone levels and the risk of anemia in older men and women". In: *Arch Intern Med* 166.13 (2006), pp. 1380–8.

[501] M. J. Koury and P. Ponka. "New insights into erythropoiesis: the roles of folate, vitamin B12, and iron". In: *Annu Rev Nutr* 24.1 (2004), pp. 105–131.

[502] J. J. Carrero et al. "Testosterone deficiency is a cause of anaemia and reduced responsiveness to erythropoiesis-stimulating agents in men with chronic kidney disease". In: *Nephrol Dial Transplant* 27.2 (2012), pp. 709–15.

[503] C. N. Roy et al. "Association of testosterone levels with anemia in older men: A controlled clinical trial". In: *JAMA Intern Med* 177.4 (2017), pp. 480–490.

[504] S. O. Schwartz and V. C. Flowers. "Morphologic changes in red blood cell with iron deficiency anemia". In: *JAMA* 130.10 (1946), pp. 622–624.

[505] B. George-Gay and K. Parker. "Understanding the complete blood count with differential". In: *J Perianesth Nurs* 18.2 (2003), pp. 96–117.

[506] E. Bachman et al. "Testosterone induces erythrocytosis via increased erythropoietin and suppressed hepcidin: evidence for a new erythropoietin/hemoglobin set point". In: *J Gerontol A Biol Sci Med Sci* 69.6 (2014), pp. 725–35.

[507] D. Furman et al. "Chronic inflammation in the etiology of disease across the life span". In: *Nat Med* 25.12 (2019), pp. 1822–1832.

[508] K. Ghoshal and M. Bhattacharyya. "Overview of platelet physiology: Its hemostatic and nonhemostatic role in disease pathogenesis". In: *Scientific World Journal* 2014 (2014), p. 781857.

[509] J. Y. Choi et al. "Iron intake, oxidative stress-related genes (MnSOD and MPO) and prostate cancer risk in CARET cohort". In: *Carcinogenesis* 29.5 (2008), pp. 964–970.

[510] Study of the Effectiveness of Additional Reductions in Cholesterol Homocysteine Collaborative Group. "Effects of homocysteine-lowering with folic acid plus vitamin B12 vs placebo on mortality and major morbidity in myocardial infarction survivors: A randomized trial". In: *JAMA* 303.24 (2010), pp. 2486–2494.

[511] K. Wu et al. "Associations between unprocessed red and processed meat, poultry, seafood and egg intake and the risk of prostate cancer: A pooled analysis of 15 prospective cohort studies". In: *Int J Cancer* 138.10 (2016), pp. 2368–82.

[512] G. Multhoff, M. Molls, and J. Radons. "Chronic inflammation in cancer development". In: *Front Immunol* 2 (2011), p. 98.

[513] L. M. Coussens and Z. Werb. "Inflammation and cancer". In: *Nature* 420 (2002), p. 860.

[514] N. Adris et al. "A prospective cohort examination of haematological parameters in relation to cancer death and incidence: the Busselton Health Study". In: *BMC Cancer* 18.1 (2018), pp. 863–863.

[515] K. H. Allin, S. E. Bojesen, and B. G. Nordestgaard. "Inflammatory biomarkers and risk of cancer in 84,000 individuals from the general population". In: *Int J Cancer* 139.7 (2016), pp. 1493–1500.

[516] M. Van Hemelrijck et al. "Association between levels of C-reactive protein and leukocytes and cancer: Three repeated measurements in the Swedish AMORIS study". In: *Cancer Epidemiol Biomarkers Prev* 20.3 (2011), pp. 428–437.

[517] S. E. R. Bailey et al. "Clinical relevance of thrombocytosis in primary care: a prospective cohort study of cancer incidence using English electronic medical records and cancer registry data". In: *Br J Gen Pract* 67.659 (2017), e405–e413.

[518] K. L. Margolis et al. "Prospective study of leukocyte count as a predictor of incident breast, colorectal, endometrial, and lung cancer and mortality in postmenopausal women". In: *Arch Intern Med* 167.17 (2007), pp. 1837–44.

[519] A. Fry et al. "Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population". In: *Am J Epidemiol* 186.9 (2017), pp. 1026–1034.

[520] UK Biobank. *Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource.* URL: `http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UKBiobank_genotyping_QC_documentation-web.pdf`.

[521] S. MacMahon 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". In: *Lancet* 335.8692 (1990), pp. 765–74.

[522] A. Hunt, D. Harrington, and S. Robinson. "Vitamin B12 deficiency". In: *BMJ* 349 (2014).

[523] C. F. Snow. "Laboratory diagnosis of vitamin b12 and folate deficiency: A guide for the primary care physician". In: *Arch Intern Med* 159.12 (1999), pp. 1289–1298.

[524] G. d'Onofrio et al. "Simultaneous measurement of reticulocyte and red blood cell indices in healthy subjects and patients with microcytic and macrocytic anemia". In: *Blood* 85.3 (1995), pp. 818–23.

[525] J. C. Barton, L. F. Bertoli, and B. E. Rothenberg. "Peripheral blood erythrocyte parameters in hemochromatosis: Evidence for increased erythrocyte hemoglobin content". In: *J Lab Clin Med* 135.1 (2000), pp. 96–104.

[526] G. Murphy et al. "Cancer risk after pernicious anemia in the US elderly population". In: *Clin Gastroenterol Hepatol* 13.13 (2015), pp. 2282–9.

[527] S. Caini et al. "Sexually transmitted infections and prostate cancer risk: a systematic review and meta-analysis". In: *Cancer Epidemiol* 38.4 (2014), pp. 329–38.

[528] A. T. Toriola et al. "Prediagnostic circulating markers of inflammation and risk of prostate cancer". In: *Int J Cancer* 133.12 (2013), pp. 2961–7.

[529] J. L. Babitt and H. Y. Lin. "Mechanisms of anemia in CKD". In: *J Am Soc Nephrol* 23.10 (2012), pp. 1631–1634.

[530] G. Wong et al. "Chronic kidney disease and the risk of cancer: an individual patient data meta-analysis of 32,057 participants from six prospective studies". In: *BMC Cancer* 16 (2016), p. 488.

[531] M. M. Edey. "Male sexual dysfunction and chronic kidney disease". In: *Front Med (Lausanne)* 4 (2017), pp. 32–32.

[532] G. Bellinghieri et al. "Sexual dysfunction in chronic renal failure". In: *J Nephrol* 21 Suppl 13 (2008), S113–7.

[533] P. Kamenický et al. "Growth hormone, insulin-like growth factor-1, and the kidney: Pathophysiological and clinical implications". In: *Endocrine Reviews* 35.2 (2014), pp. 234–281.

[534] R. Pecoits-Filho et al. "Interactions between kidney disease and diabetes: dangerous liaisons". In: *Diabetol Metab Syndr* 8.1 (2016), p. 50.

[535] J. L. Spivak. "The anaemia of cancer: death by a thousand cuts". In: *Nat Rev Cancer* 5.7 (2005), pp. 543–55.

[536] D. G. Menter et al. "Platelets and cancer: a casual or causal relationship: revisited". In: *Cancer Metastasis Rev* 33.1 (2014), pp. 231–269.

[537] A. Swierczak et al. "Neutrophils: important contributors to tumor progression and metastasis". In: *Cancer Metastasis Rev* 34.4 (2015), pp. 735–751.

[538] S.B. Coffelt, M.D. Wellenstein, and K.E. de Visser. "Neutrophils in cancer: neutral no more". In: *Nat Rev Cancer* 16 (2016), p. 431.

[539] S. J. Gandhi et al. "Prevalence, comorbidity and investigation of anemia in the primary care office". In: *J Clin Med Res* 9.12 (2017), pp. 970–980.

[540] G.C. Kabat et al. "White blood cell count and total and cause-specific mortality in the Women's Health Initiative". In: *Am J Epidemiol* 186.1 (2017), pp. 63–72.

[541] P. J. Vinholt et al. "Platelet count is associated with cardiovascular disease, cancer and mortality: A population-based cohort study". In: *Thromb Res* 148 (2016), pp. 136–142.

[542] V. Orrù et al. "Genetic variants regulating immune cell levels in health and disease". In: *Cell* 155.1 (2013), pp. 242–256.

[543] C. J. D. Wallis, Y. Krakowsky, and A. Finelli. "Low testosterone and prostate cancer: is the protection real?" In: *Eur Urol* 74.5 (2018), pp. 595–596.

[544] S. J. Otto, F. H. Schröder, and H. J. De Koning. "Low all-cause mortality in the volunteer-based Rotterdam section of the European randomised study of screening for prostate cancer: self-selection bias?" In: *J Med Screen* 11.2 (2004), pp. 89–92.

[545] P. F. Pinsky et al. "Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial". In: *Am J Epidemiol* 165.8 (2007), pp. 874–881.

[546] M. R. Munafo et al. "Collider scope: when selection bias can substantially influence observed associations". In: *Int J Epidemiol* 47.1 (2018), pp. 226–235.

[547] L. Richiardi et al. "Baseline selection on a collider: a ubiquitous mechanism occurring in both representative and selected cohort studies". In: *J Epidemiol Community Health* 73.5 (2019), pp. 475–480.

[548] G. D. Batty et al. "Generalisability of results from UK Biobank: Comparison with a pooling of 18 cohort studies". In: *medRxiv* (2019), p. 19004705.

[549] H. M. Blanck et al. "Laboratory issues: use of nutritional biomarkers". In: *J Nutr* 133 Suppl 3 (2003), 888s–894s.

[550] E. White. "Measurement error in biomarkers: sources, assessment, and impact on studies". In: *IARC Sci Publ* 163 (2011), pp. 143–61.

[551] U. Srinivas-Shankar et al. "Effects of testosterone on muscle strength, physical function, body composition, and quality of life in intermediate-frail and frail elderly men: a randomized, double-blind, placebo-controlled study". In: *J Clin Endocrinol Metab* 95.2 (2010), pp. 639–650.

[552] M. B. Cook et al. "Relationships between circulating and intraprostatic sex steroid hormone concentrations". In: *Cancer Epidemiol Biomarkers Prev* 26.11 (2017), pp. 1660–1666.

[553] J. A. C. Sterne et al. "Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls". In: *BMJ* 338 (2009), b2393–b2393.

[554] V. Zarotsky et al. "Systematic literature review of the risk factors, comorbidities, and consequences of hypogonadism in men". In: *Andrology* 2.6 (2014), pp. 819–34.

[555] P. J. Goodman et al. "The Prostate Cancer Prevention Trial: Design, biases and interpretation of study results". In: *J Urol* 175.6 (2006), pp. 2234–2242.

[556] P.J. Goodman et al. "Long-term effects of finasteride on prostate cancer mortality". In: *N Engl J Med* 380.4 (2019), pp. 393–394.

[557] P. Sedgwick. "Retrospective cohort studies: advantages and disadvantages". In: *BMJ* 348 (2014), g1072.

[558] P. Sedgwick. "Prospective cohort studies: advantages and disadvantages". In: *BMJ* 347 (2013), f6726.

[559] T. Hill et al. "Data linkage reduces loss to follow-up in an observational HIV cohort study". In: *J Clin Epidemiol* 63.10 (2010), pp. 1101–1109.

[560] W. C. Allsbrook Jr. et al. "Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologist". In: *Hum Pathol* 32.1 (2001), pp. 81–8.

[561] D. A. Lawlor, K. Tilling, and G. Davey Smith. "Triangulation in aetiological epidemiology". In: *Int J Epidemiol* 45.6 (2016), pp. 1866–1886.

[562] J. Yarmolinsky et al. "Causal inference in cancer epidemiology: what is the role of Mendelian randomization?" In: *Cancer Epidemiol Biomarkers Prev* 27.9 (2018), pp. 995–1010.

[563] PRACTICAL. *The Institute of Cancer Research PRACTICAL: Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome*. URL: `http://practical.icr.ac.uk/blog/`.

[564] B. L. Pierce, H. Ahsan, and T. J. VanderWeele. "Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants". In: *Int J Epidemiol* 40.3 (2010), pp. 740–752.

[565] J. R. Staley and S. Burgess. "Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization". In: *Genet Epidemiol* 41.4 (2017), pp. 341–352.

[566] R. J. Silverwood et al. "Testing for non-linear causal effects using a binary genotype in a Mendelian randomization study: application to alcohol and cardiovascular traits". In: *Int J Epidemiol* 43.6 (2014), pp. 1781–90.

[567] L. Egevad et al. "Prognostic value of the Gleason score in prostate cancer". In: *BJU Int* 89.6 (2002), pp. 538–42.